IJCSIS Vol. 10 No. 1, January 2012
ISSN 1947-5500

International Journal of Computer Science & Information Security

Editorial
Message from Managing Editor

The IJCSIS continues to be a leading scholarly journal in computer science, networks, security and emerging technologies. To a large extend, the credit for high quality, visibility and recognition of the journal goes to the editorial board and the technical review committee.

The journal covers the frontier issues in Information and Communication technology, and computer science and their applications in business, industry and other subjects. (See monthly Call for Papers)

For complete details about IJCSIS archives publications, abstracting/indexing, editorial board and other important information, please refer to IJCSIS homepage. IJCSIS appreciates all the insights and advice from authors/readers and reviewers.

We look forward to receive your valuable papers. If you have further questions please do not hesitate to contact us at ijcsiseditor@gmail.com. Our team is committed to provide a quick and supportive service throughout the publication process.

A complete list of journals can be found at:

http://sites.google.com/site/ijcsis/

IJCSIS Vol. 10, No. 1, January 2012 Edition

ISSN 1947-5500 © IJCSIS, USA.

Journal Indexed by (among others):

IJCSIS EDITORIAL BOARD

TABLE OF CONTENTS

Adaptive Optical PIC Applied in VLC For Multi-user Access Interference Reduction

Peixin, Li
Department of Electronics and Radio Engineering
Kyung Hee University
Suwon, Korea
peixin@khu.ac.kr

Ying Yi
Department of Electronics and Radio Engineering
Kyung Hee University
Suwon, Korea
yiying@khu.ac.kr

***Abstract*—Optical wireless data transmission systems for indoor application are usually affected by optical interference induced by sun light and artificial ambient lights. This paper presents a characterization of the optical interference produced in visible light communication (VLC) systems and proposes an effective scheme to solve it. Regarding the sun light noise and some artificial light noises reduction, the common method is to adapt the optical bandpass filter which can distinguish the wavelength between interference lights and information lights. However, for some photo-electric systems, the visible lights from the transmitters occupy the same wavelength range, in this case, the optical bandpass filter would not reduce the interference noise from the other user, for example, the optical interference caused by multi-user access of the optical medium. Therefore, we proposed a novel scheme, adaptive optical parallel interference cancellation (AOPIC) to reduce the multiple access interference (MAI) and multiple user interference (MUI) induced by multi-user access of the optical medium, the conventional parallel interference cancellation (PIC) is analyzed as the comparison. Through the simulation results, we can conclude that the AOPIC scheme shows much better bit error rate (BER) performance than the conventional PIC with the increasing number of user.**

Keywords-component; AOPIC; MAI; MUI;MC-CDMA

I. INTRODUCTION

Recently, visible light communication (VLC) systems have attracted attentions due to the growing progress in the field of visible light technology [1]. Visible light has several attractive features distinct from those of radio frequency (RF) and infrared (IR) [2]. Though both LED and laser diodes (LD) are usually used as optical sources, LEDs are preferred as strong candidates for the next generation lighting technology [3] for several reasons including fewer safety concerns, a relatively long useful life time, and a wider emission angle than those of LDs [4]. As an emitter for optical wireless communication, LED lights emit visible rays as the medium of optical data transmission. Nevertheless, with the development of VLC systems, both the industrial and scientific communities have recognized that visible light also can be used in the high data rate transmission systems, since it has the following advantages compared to those of RF:

- VLC is harmless to our health.
- A friendly user interface.
- A lighting device is used as a transmitter without any traces of embellishment in the wireless communication environment.
- The visible light spectrum does not occupy the radio frequency spectrum; therefore the electromagnetic interference (EMI) can be avoided by VLC.

VLC is suitable for high-speed data transmission, especially in an indoor environment. Though VLC system has distinct advantages as mentioned above, the performance of VLC is limited by several aspects, for example, an inevitable issue is the optical interference noise that induced by both natural and artificial light on the receiving photodiode (PD) and the optical interference from the multi-user access of the optical medium. In addition, few studies have examined the effects of optical interference from the multi-user access of the optical medium. Actually, in the realistic communication environment, numerous users transmitting signals are inevitable existing, Motivated by this, we investigate the utilization of Multi-Carrier (MC)-code division multiple access (CDMA) technology in VLC system. However, we also found that the interference signals from the other users will cause the multi-user interference (MUI) and multi-access interference (MAI) that have a significant impact on the MC-CDMA communication performance. We further proposed an effective scheme, AOPIC, to reduce both MAI and MUI induced by multi-user access of the optical medium, the conventional parallel interference cancellation (PIC) is analyzed as the comparison. Our essential target is to improve the end-to-end optical wireless communication performance.

The remainder of this paper is organized as follows. In Section Ⅱ, the solution scheme for the natural light noise is described, and the performance comparisons and analyses are given for the proposed system using the AOPIC technique and a typical PIC technique through computer simulations in Section Ⅲ. Section Ⅳ provides the concluding remarks.

II. NATURAL LIGHT NOISE REDUCTION

A. Sunlight interference

Sunlight that produces interference to the desired lights is the dominant noise source to induce power penalties in the performance of transmission systems [5-8], and usually this

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

power penalties are very large. The effects of optical interference have been included in the performance analysis of VLC high data rate transmission systems by considering that the optical interference power is proportional to the average surface area on the PD. Actually, the sun light produces the highest levels of power spectral density around the wavelength area of visible light, therefore, it is the major source of optical interference on the receiver PD. Moreover, the wavelength of the other artificial ambient lights (such as, incandescent lights and fluorescent lights) and the wavelength of the transmitted visible light overlap in some area. Thereby shot noise and interference are induced.

The optical interference induces a power penalty that in some cases may be very large. Therefore, optical filtering is used in most systems to overcome some of the problems produced by the ambient light interference. The higher efficiency of the optical filter is achieved for sunlight interference reduction due to the differences in the optical spectrum of each light source [9-10]. Usually, optical filter includes two types, long-pass filter and band-pass filter (or interference filters). The use of optical filters reduces the amount of ambient light that reaches the PD, thus reducing the undesirable effects. The transmission gain obtained by the use of an optical filter depends on its efficiency in attenuating the ambient light while keeping intact the transmitted signal. Clearly, interference (band-pass) optical filters are more efficient in that operation, provided that the transmitted signal is not attenuated as well.

B. PIN PD Receiver

The front-end of PIN PD receiver is constructed from an optical bandpass filter, a concentrator, a positive-intrinsic-negative (PIN) silicon PD, and a preamplifier. Optical filter could reduce the optical interference without the bandpass wavelength at visible light spectrum but it could not eliminate the interference light that are over the same wavelength as desired light. The total fraction of power transmitted through the filter, assuming lossless dielectrics, is given by:

$$T(\theta_1) = 1 - \frac{1}{2}\left(\left|\rho_{TE}\right|^2 + \left|\rho_{TM}\right|^2\right) \tag{1}$$

where the reflection coefficients ρ_{TE} and ρ_{TM} are defined by the following set of recursive equations [11]—[13]

$$\rho = \frac{N_1 - \eta_2}{N_1 + \eta_2} \tag{2}$$

$$N_k = \begin{cases} n_k / \cos\theta_k, & for\ TE \\ n_k \cos\theta_k, & for\ TM \end{cases}, \quad k \in \{2,...,K\} \tag{3}$$

$$\eta_k = N_k \frac{\eta_{k+1}\cos\beta_k + jN_k \sin\beta_k}{N_k \cos\beta_k + j\eta_{k+1}\sin\beta_k}, \; k \in \{2,...,K\} \tag{4}$$

$$\theta_k = \sin^{-1}\left(\frac{n_{k-1}}{n_k}\sin\theta_{k-1}\right), \quad k \in \{2,...,K\} \tag{5}$$

Here, θ_k is the angle made by the light ray as it passes from medium k to medium k + 1, η_k is the effective complex valued index "seen" by the light wave as it enters medium k, and $\beta_k = 2p\cos(\theta_k)n_k d_k / \lambda$, where λ is the wavelength of the light in a vacuum [11]. Starting with $\eta_k = N_k$, (4) can be applied recursively to arrive at η_2, which when substituted into (2) yields ρ_{TE} or ρ_{TM}, depending on the initialization of the $\{N_k\}$ as either TE or TM in (3).

A PIN silicon PD is suitable for the outdoor environment because of its fast switching capability. Regarding the preamplifier, we design a low noise field-effect-transistor (FET)-based transimpedance inside [14]. So the total received noise variance is the sum of contributions from the shot noise and thermal noise, given by [14]:

$$\begin{aligned} \sigma_{total}^2 &= \sigma_{shot}^2 + \sigma_{thermal}^2 \\ &= 2qg\gamma P_{bg} I_2 B + \left\{\frac{8\pi k T_k}{g} \eta A I_2 B^2 \right. \\ &\quad \left. + \frac{16\pi^2 k T_k \Gamma}{g_m} \eta^2 A^2 I_3 B^3\right\} \end{aligned} \tag{6}$$

where the first term is shot noise, and second term is thermal noise variance. q is the electronic charge (1.6×10^{-19} C), and B is the equivalent noise bandwidth corresponding to the data rate. γ is the O/E conversion efficiency, and P_{bg} is the optical power of the background light, which varies with time and reaches its peak at noon. T_k is the absolute temperature of the environment, g is the open-loop voltage gain, η is the fixed capacitance of the PD per unit area, A is the physical area of the PD, Γ is the FET channel noise factor, g_m is the FET transconductance, k is Boltzmann's constant. We defined the noise bandwidth factor I_2=0.562 following [15-16], and the noise bandwidth factor I_3=0.0868. We choose the average temperature and background noise power according to the time of day from [17], and choose the other parameter values of the referred symbols from [15-16] and list them in Table Ⅰ.

Sun light produces interference due to the time variations on its intensity as shown in Fig. 1. In the simulation as depicted in Fig. 1, the proposed receiver including PIN PD can significantly reduce the optical interference from the sun light. When the transmitted date rate is 10 Mbps, the PIN receiver can reduce nearly 10 dBm interference power as compared to the original sunlight power without handpass filter. Therefore the simulation result as shown in Fig. 1 can illustrates that optical handpass filter is effectively used to reduce the optical interference noise with the different wavelength to desired lights.

TABLE I. PARAMETERS FOR OPTICAL INTERFERENCE REDUCTION CALCULATION.

open-loop voltage gain, g	10
fixed capacitance, η	112 [pF/cm^2]
FET transconductance, g_m	30 [mS]
Noise bandwidth factor, I_2	I_2=0.562
Noise bandwidth factor, I_3	I_3 =0.0868
FET channel noise factor, Γ	1.5
Data rate, R_b	10, 50 [Mbit/s]

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

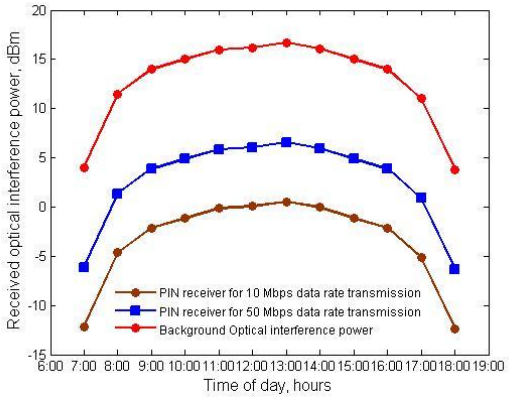


Figure 1. Optical power of interference noise over the day and the reduction effects by the proposed receiver for different transmission data rate.

III. PROPOSED SYSTEM MODEL

A. MC-CDMA

Different from the frequency division multiple access (FDMA) and time division multiple access (TDMA), conventional CDMA techniques use spread codes to identify each user separately, however, all users in a CDMA system interfere with each other. Take an example of MC-CDMA, MC-CDMA is a combination access techniques of CDMA and orthogonal frequency division multiplexing (OFDM) [18]. Regarding the signals from the other user, it is always considered as noise, for example, MAI and MUI. These interferences cause communication performance degradation and limit the capacity of CDMA systems. Conventional CDMA systems independently detect each user in parallel using a matched filter which consists of the unique spreading code used by that user. In the MC-CDMA, the transmitted signal of the *k*-th user is given by:

$$s_k(t) = \sum_{m=1}^{M} \sqrt{2P_{k,m}}\, b_{k,m}(t)\, c_k(t) \cos(2\pi f_m t + \theta_{k,m}) \qquad (7)$$

where M is the total number of sub-carriers, $P_{k,m}$ represents the transmitted power over *m*-th sub-carrier for the *k*-th user. The subcarriers in MC-CDMA are orthogonal over the chip duration, hence, *m*-th sub-carrier frequency is $f_m=f_0+m/T_c$, where T_c is chip duration. $\theta_{k,m}$ is the phase angle introduced in the carrier modulation process which distributes over $[0, 2\pi]$. $b_{k,m}(t)$ and $c_k(t)$ are the data sequence and spreading waveform, respectively, given as follows:

$$\begin{aligned} b_{k,m}(t) &= \sum_{i=-\infty}^{\infty} b_{k,m} \times \Pi_b(t - iT_s) \\ c_k(t) &= \sum_{i=-\infty}^{\infty} c_k \times \Pi_c(t - iT_c) \end{aligned} \qquad (8)$$

where $b_{k,m}$ and c_k are independent random variables with equal probability of +1 or -1. While Π_b is the rectangular symbol waveform that is defined over the symbol duration T_s, and Π_c is the rectangular chip waveform over the interval $[0, T_c]$.

We consider K asynchronous MC-CDMA users, all of whom have the same number of subcarriers M and the same spreading factor. Because the carrier frequency of visible light is very high, the multipath fading can be ignored in the optical channels [19]. Consequently, the received signals can be written as:

$$\begin{aligned} r(t) = \sum_{k=1}^{K} \sum_{m=1}^{M} \Big\{ &\sqrt{2\tilde{P}_{k,m}}\, b_{k,m}(t - \tau_k) c_k(t - \tau_k) \\ &\cos(2\pi f_m t + \varphi_{k,m}) \Big\} + n(t) \end{aligned} \qquad (9)$$

where n(t) is the additive white Gaussian noise (AWGN). τ_k is the time delay for the *k*-th user. $\varphi_{k,m}$ express the uniform random variables over $[0, 2\pi]$. $\tilde{P}_{k,m}$ is the received power, the relationship between received power and transmitted power are given by:

$$\tilde{P}_{k,m} = P_{k,m}\left(\frac{n+1}{2\pi d^2}\right) A \cos^n \phi T_s(\psi) G \cos\psi \qquad (10)$$

Where *A* is the physical area of PD, *d* is the distance between the emitter and the receiver, $T_s(\Psi)$ is the gain of the optical filter, Ψ is the angle of incidence, and G is the optical concentrator gain [20], as shown as follows:

$$G = \frac{n^2}{\sin^2 \psi_c} \qquad (11)$$

where n is the material refractive index and Ψ_c denotes half of the concentrator FOV, usually $\Psi c \le \pi/2$.

The sampled output of the match filter for the k-user in typical MC-CDMA systems can be expressed as follows:

$$\begin{aligned} \hat{r}_k(t) &= \int_0^{Tc} r(t) c_k(t) \cos(2\pi f_m t + \varphi_{k,m}) dt \\ &= \int_0^{Tc} c_k(t) \cos(2\pi f_m t + \varphi_{k,m}) \\ &\quad \cdot [\sum_{k=1}^{K} \sum_{m=1}^{M} \sqrt{2\tilde{P}_{k,m}}\, b_{k,m}(t - \tau_k) \\ &\quad \cdot c_k(t - \tau_k) \cos(2\pi f_m t + \varphi_{k,m}) + n(t)] dt \end{aligned} \qquad (12)$$

If the time delay is limited in a small value, the (12) can be written as:

$$\begin{aligned} \hat{r}_k(t) = &\sum_{m=1}^{M} \sqrt{2\tilde{P}_{k,m}}\, b_{k,m}(t) + \sum_{\substack{j=1 \\ j \neq k}}^{K} \sum_{m=1}^{M} \sqrt{2\tilde{P}_{j,m}}\, b_{j,m}(t) \rho_{kj} \\ &+ \int_0^{Tc} c_k(t) n(t) \cos(2\pi f_m t + \varphi_{k,m}) dt \end{aligned} \qquad (13)$$

$\hat{r}_k(t)$ consists of three terms. The first is the desired signal which gives the sign of the information bit b_k. The second term is the result of the MAI, and the last is due to noise. The cross-correlation of the spreading codes between k-user and j-user is:

$$\rho_{kj} = \int_0^{Tc} c_k(t) c_j(t) dt \qquad (14)$$

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

The decision made by the conventional single-user receiver is given as:

$$b_k = sign[\hat{r}_k(t)] \tag{15}$$

where sign [.] is the sign function. Hence, the single-user matched filter receiver takes the MAI as noise and it can't suppress MAI. So we have to propose the interference cancellation scheme to further reduce the MAI.

B. AOPIC

The conventional PIC detector cancels the estimates of the MAI from the outputs of the matched filters in a parallel manner. It follows an iterative process. Thus,

$$b_k^{z+1} = sign[r_k - \sum_{j \neq k}^{K} \sqrt{2P_j} b_j^{z+1} \rho_{kj}] \tag{16}$$

PIC detects all users simultaneously, and parallel detection can be repeated. This process can be repeated over several stages. With the increase of the stage in PIC process, the better BER performance can be obtained, but at the cost of high complexity. On the contrary, AOPIC is based on mean square error (MSE) criteria, the cost function is given as follows:

$$\min_{W} E[\left| r(t) - \hat{r}_{(z)}(t) \right|^2] \tag{17}$$

Where r(t) is defined in (9), W is the weight vector. $\hat{r}_{(z)}(t)$ represents the estimate of the received signal at the *z*-th sequence of iterations that is defined as follows:

$$\hat{r}_{(z)}(t) = \sum_{k=1}^{K} \sum_{m=1}^{M} \hat{b}_{k,m}^{(z)} c_k^{(z)} \cos(2\pi f_m t + \varphi_{k,m}) W_{k,m}^{(z)} \tag{18}$$

$\hat{b}_{k,m}^{(z)}$ is the estimate of $b_{k,m}$ at the (*z*)-th iteration.

The proposed system is depicted by Fig. 2. Frequency mapping accomplishes data transmission within the visible light wavelength range. Intensity modulation (IM) and photo-detector (PD) complete the conversion between the electrical signal and the optical signals. Spread codes are used to distinguish different users' data, since users' data are separated on the basis of their signature waveforms. The entire concept of AOPIC is based on the premise that the received signal can be reliably estimated. Decision, as shown in Fig. 2, follows an iterative process and subtracts the interference from other users. Channel estimation evaluates all users simultaneously and then AOPIC can be repeated to update the weight vector. Many algorithms can effectively reduce the MSE, for example, least mean square (LMS) and recursive least square (RLS). Since the LMS algorithm has a slower complexity as compared to the RLS, we propose a LMS algorithm in the AOPIC approach. The input of the first stage of AOPIC is defined as:

$$\hat{b}_k^{(z)} = sign[\hat{r}_k(t)] \tag{19}$$

where $\hat{r}_k(t)$ is defined in (13). The optimum weights are derived via a LMS algorithm which operates as follows:

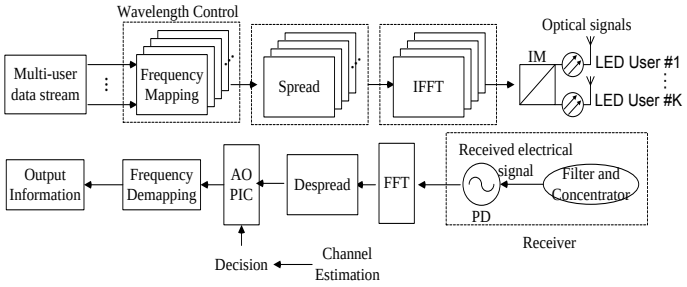


Figure 2. Simplified block diagram of proposed system model.

$$W_{k,m}^{(z+1)} = W_{k,m}^{(z)} + \frac{\alpha}{\left\| \hat{s}_{k,m}^{(z)} \right\|^2} \hat{s}_{k,m}^{(z)} [e^{(z)}]^* \tag{20}$$

where α is a step size, $\hat{s}_{k,m}^{(z)}$ denotes the input vector of the LMS equalizer, and it is defined as:

$$\hat{s}_{k,m}^{(z)} = \hat{b}_{k,m}^{(z)} c_k^{(z)} \tag{21}$$

And * denotes complex conjugate, e(z) is the error between the desired response and the output of the LMS filter, so that,

$$e^{(z)} = r - \hat{r}_{(z)} \tag{22}$$

Weight vector $W_{k,m}^{(z)}$ is updated iteratively via minimize the e given as follows:

$$e^{(z)} = \int_0^{Tc} \left| r(t) - \sum_{k=1}^{K} \sum_{m=1}^{M} \hat{b}_{k,m}^{(z)} c_k^{(z)} \cos(2\pi f_m t + \varphi_{k,m}) W_{k,m}^{(z)} \right|^2 dt \tag{23}$$

Consider the k-th user, the interference cancellation can be performed as:

$$\hat{r}_k^{(z)}(t) = r^{(z)}(t) - \sum_{\substack{j=1 \\ j \neq k}}^{K} \sum_{m=1}^{M} \hat{s}_{k,m}^{(z)} \cos(2\pi f_m t + \varphi_{k,m}) W_{k,m}^{(z)} \tag{24}$$

Therefore, the decision $\hat{b}_{k,m}^{(z)}$ in (21) becomes more reliable, since it is based on the less interfered signal $\hat{r}_k^{(z)}$.

TABLE II. PARAMETERS FOR MULTI-USERS BER CALCULATION.

Modulation:	BPSK
Noise Model:	AWGN
Spread code:	Walsh code
IFFT/FFT Size:	64
Number of users:	5, 16, 48
Spread factor:	32
Original Data rate:	100 Mbps
O/E Conv. Efficiency:	0.53 [A/W]
Background Light Noise:	0 [dBm]

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

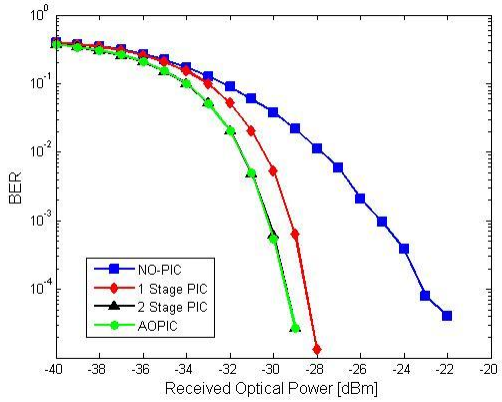


Figure 3. Comparison of BER of AOPIC, conventional PIC and no-PIC scheme versus SNR for 5 users.

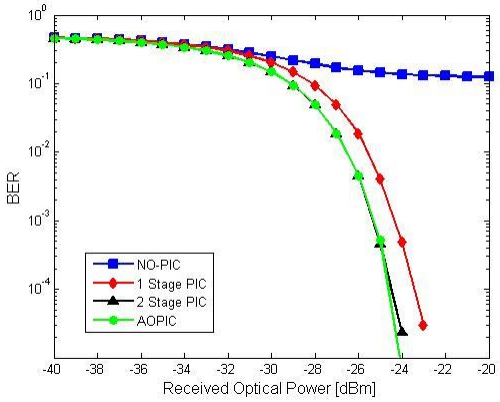


Figure 5. Comparison of BER of AOPIC, conventional PIC and no-PIC scheme versus SNR for 48 users.

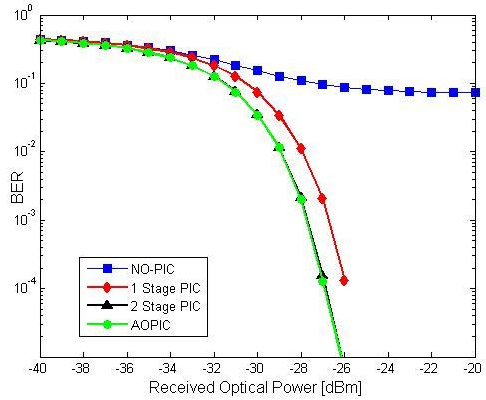


Figure 4. Comparison of BER of AOPIC, conventional PIC and no-PIC scheme versus SNR for 16 users.

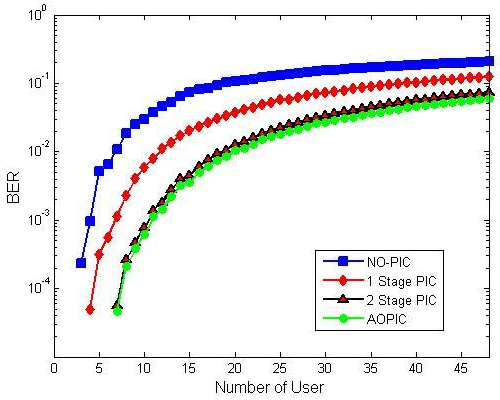


Figure 6. Comparison of BER of AOPIC, conventional PIC and no-PIC scheme versus the number of user for SNR=18dB.

C. Simulation Analysis

As analyzed above, we adopted Walsh spread code in the AOPIC scheme and chose multi-stage PIC scheme for comparison purposes. The received electrical signal-to-noise ratio (SNR) from [12] is:

$$SNR = \frac{(\gamma P_r)^2}{\sigma_{total}^2} \tag{25}$$

where γ is the O/E conversion efficiency. σ_{total} is defined in (6). P_r represents the received optical power. The simulation parameters are listed in Table Ⅱ. Based on Table 1 and 2, the simulation results are given in Figs 3-6.

It is shown in Figs. 3-6 that significant performance improvement was obtained by employing the AOPIC and the 2-stage cancellation into the PIC receiver. It is clear from Figs. 3-6 that the 2-stage PIC shows a significantly better performance than the 1-stage PIC regardless of the number of users. As a comparison, 2-stage PIC obtains the better BER performance, however, at a cost of high complexity. Our proposed AOPIC scheme is much easier operated, and from Figs. 3-6, we can find that AOPIC scheme shows a closed BER performance as compared to the 2-stage PIC.

In Fig. 3, the BER performance penalty can be compensated with the increase of received optical power in both PIC scheme and AOPIC scheme if the number of user is small. However, with the increase of the number of user, the degree of MUI and MAI caused by the multiple users becomes larger, as shown in Figs. 4-5, the scheme without PIC has been totally failed. Though the PIC can compensate the BER performance penalty with the SNR increases, we can find that AOPIC and 2-stage PIC are shown to achieve better performance than the conventional PIC. AOPIC shows a similar BER performance to 2-stage PIC. When the number of users becomes much larger, 48 users, as shown in Fig. 5, AOPIC provides a relative better performance than does 2-stage PIC.

Finally, we assume that the SNR at the receiver is 18dB, and from Fig. 6, we can further observe that the AOPIC shows an excellent performance among all schemes with an increase in the number of users. Therefore, we can conclude that the Walsh code has a better orthogonal quality in distinguishing different users' data, and AOPIC can further suppress the MUI and MAI effectively. It is shown that AOPIC can retain a performance advantage over conventional PIC.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

IV. CONCLUSIONS

The performance of visible light data transmission systems for indoor use is severely impaired by the optical interference noise induced by natural and artificial ambient light. In order to combat the effects of ambient light on the system performance, optical filtering is usually adopted. However, even when resorting to optical filter, the optical noise penalty imposed by the interference from the other user may be difficult to be compensated. In particular, with the increasing number of users, the MUI and MAI induced by multi-user access of the optical medium imposes very large performance penalties on systems operating at data rates up to a few tens of Mbps.

In this paper, a conventional technique to overcome the penalty induced by ambient light interference is analyzed. This technique explores the different optical wavelength of the transmitted signal and the ambient interference light and the characteristics of optical bandpass filtering to cancel the interfering signal. Some aspects of its implementation are also discussed. Moreover, it is well known that the MAI and MUI limit MC-CDMA system capacity and reduce communication performance. Therefore, we also present an AOPIC scheme for a MC-CDMA system, using band-limited spreading waveforms to prevent the MAI and MUI. The AOPIC receiver parallel detects the interferers' signals and subtracts them from the user-of-interest. A comparison is made among conventional PIC, 2-stage PIC, AOPIC. The results obtained with AOPIC are shown to be much better than those obtained through the other interference cancellation schemes.

REFERENCES

[1] D.C.O'Brien et al, " Visible lightcommunication: state of the art and prospects," published in Proc. Wireless World Research Forum 2007.

[2] M.Z.Afgani, H.Haas, H.Elgala, D.Knipp, "Visible light communication using OFDM," Proc. IEEE Symp. on Wireless Pervasive Computing, TRIDENTCOM 2006.

[3] C.P.Kno, R. M. Fletcher, T. D. Owentowski, M.C.Lardizabal and M.G.Craford, "High performance ALGaInP visible light-emitting diodes," Appl. Phys. Lett., vol. 57, no.27, pp. 2937-2939, 1990.

[4] K.D.Langer and J.Grubor, "Recent Developments in Optical Wireless Communications using Infrared and Visible Light", ICTON, 2007, pp.146-151.

[5] A.M.R. Tavares, A.J.C. Moreira, C. Lomba, L. Moreira, R.T. Valadas and A.M. de Oliveira Duarte, Experimental results of a 1 Mbps IR transceiver for indoor wireless local area networks, in: COMCON V-Intern. Conf. on Advances in Communications & Control, Crete, Greece (June 26–30, 1995).

[6] R.T. Valadas, A.J.C. Moreira, C. Oliveira, L. Moreira, C. Lomba, A.M.R. Tavares and A.M. de Oliveira Duarte, Experimental results of a pulse position modulation infrared transceiver, in: Proceedings of the Seventh IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC '96), Taipei, Taiwan (October 15–18, 1996).

[7] M.D. Audeh and J.M. Kahn, Performance evaluation of baseband OOK for wireless indoor infrared LAN's operating ant 100 Mb/s, IEEE Trans. Comm. 43(6) (1995) 2085–2094.

[8] G.W. Marsh and J.M. Kahn, 50-Mb/s diffuse infrared free-space link using on-off keying with decision-feedback equalization, in: Proceedings of the Fifth IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC '94), The Hague, The Netherlands (September 1994) pp. 1086–1089.

[9] C.J. Georgopoulos, Suppressing background-light interference in an in-house infrared communication system by optical filtering, Internat. J. Optoelectronics 3(3) (1988).

[10] F.R. Gfeller and U. Bapst, Wireless in-house data communication via diffuse infrared radiation, Proc. IEEE 67(11) (November 1979).

[11] S. Ramo, J. R. Whinnery, and T. Van Duzer, "Fields and Waves in Communication Electronics" (Wiley, New York, 1984), Chap. 6, pp. 309–310.

[12] H. A. Macleod, "Thin-Film Optical Filters" (Hilger, London, 1969).

[13] J. D. Rancourt, "Optical Thin Films" (Macmillan, New York, 1987).

[14] S.D.Personick, "Receiver design for digital fiber optic communications systems, I and 11", Bell System Technical J. vol.52, no.6, pp. 843–886, July-August 1973.

[15] J.R.Barry, "Wireless infrared communications," Kluwer Academic Press, Boston, MA, 1994.

[16] A.P.Tang, J.M.Khan, and K.P.Ho, "Wireless Infrared Communication Links Using Multi-Beam Transmitters and Imaging Receivers," IEEE Int. Conf. on Communications, pp. 180–186, Dallas, TX, June 1996.

[17] I.E. Lee, M.L. Sim and F.W.L. Kung, "Performance enhancement of outdoor visible-light communication system using selective combining receiver", IET Optoelectron., Vol. 3, Iss. 1, pp. 30-39, 2009.

[18] S. Hara and R. Prasad, "Overview of multi-carrier CDMA," IEEE Com. Mag., Vol. 35, pp. 126-133, Dec.1997.

[19] Y.Tanaka, T.Komine, S.Haruyama, M. Nakagawa, "Indoor visible light data transmission system utilizing white LED lights", IEICE TRANS. COMMUN, vol.E86B, NO.8, 2003.

[20] X. Ning, R. Winston, and J. O'Gallagher, "Dielectric totally internally reflecting concentrators," Appl. Optics, vol. 26, no. 2, pp. 300–305, Jan. 1987.

AUTHORS PROFILE

Peixin Li received bachelor degree in College of Materials Science and Engineering from Jiamusi University, in Heilongjiang Province, China. He is currently pursuing the Master degree of Engineering in Department of Electronics and Radio Engineering, Kyung Hee University, Korea. His current research interests are visible light communication, MIMO-OFDM and MC/DS CDMA.

Ying Yi received the B.S degree in Information Technology from HeBei Normal University, in HeBei Province, China, and M.E. degrees from the Department of Electronics and Radio Engineering, Kyung Hee University, Korea, in 2008 and 2010, respectively. Currently, he is a research associate in Department of Electronics and Radio Engineering, Kyung Hee University, Korea. Meanwhile, he is doing the projects for IT Research and Development Program of the Korean Ministry of Knowledge Economy and Korea Evaluation Institute of Industrial Technology (MKE/KEIT) as a researcher. His research interests are optical wireless communication systems, Ad-hoc/Mesh network, and LTE.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

Performance Assessment of Tools of the Intrusion Detection/Prevention Systems

Yousef FARHAOUI
LabSiv, Equipe ESCAM
Faculty of sciences Ibn Zohr University B.P 80060, City Dakhla, Agadir, Morocco.
youseffarhaoui@gmail.com

Ahmed ASIMI
LabSiv, Equipe ESCAM
Faculty of sciences Ibn Zohr University B.P 80060, City Dakhla, Agadir, Morocco.
asimiahmed2008@gmail.com

***Abstract*— This article aims at providing (i) a general presentation of the techniques and types of the intrusion detection and prevention systems, (ii) an in-depth description of the evaluation, comparison and classification features of the IDS and the IPS and (iii) the implications of such study on how to determinate the features of some more effective IDS and IPS in the commercial domains and open source.**

***Keywords*—Intrusion Detection, Intrusion Prevention, Characteristic, Tools.**

I. INTRODUCTION

The systems of detection and prevention of intrusion, IDS and IPS, are among the most recent tools of security. According to their features, we can classify them in different kinds, for example, their techniques of detection and prevention, their architecture or the range of detection [3]. In spite of their utility, in practice most IDS/IPS experience two problems: the important number of false positives and false negatives. The false positives, the false alerts, are generated when the IDS/IPS identifies normal activities as intrusions, whereas the false negatives correspond to the attacks or intrusions that are not detected, and then no alert is generated [4]. The IDS/IPS inventors try to surmount these limitations by developing new algorithms and architectures.

Therefore, it is important for them to value the improvements brought by these new devices. In the same way, for the network and systems administrators, it would be interesting to assess the IDS/IPS to be able to choose the best before installing it on their networks or systems, but also to continue to evaluate its efficiency in operational method. Unfortunately, many false positives and false negatives persist in the new versions of the IDS/IPS, then, the brought improvements are not worthy of the continuous efforts of research and development in the domain of the detection and the prevention of intrusion. In general, it is essentially due to the absence of efficient methods of assessment of the security tools, and of the IDS/IPS in particular.

II. INTRUSION DETECTION SYSTEMS

The IDS is a mechanism which watches over the traffic network in a sneaky manner in order to mark abnormal or suspected activities and permitting to have an action of prevention on the risks of intrusions.

Mainly, there are three important distinct families of IDS:

- The NIDS, Network Based Intrusion Detection System which assures the security in the network.
- The HIDS, Host Based Intrusion Detection System which assures the security in the hosts.
- The hybrid IDS. An IDS hybrid is a combination of both the HIDS and the NIDS.

A. Network Intrusion Detection System

The NIDS are also called passive IDS since this kind of systems inform the administrator system that an attack has or had taken place, and it takes the adequate measures to assure the security of the system. The aim is to inform about an intrusion in order to look for the IDS capable to react in the post. Report of the damages is not sufficient. It is necessary that the IDS react and to be able to block the detected doubtful traffics. These reaction techniques imply the active IDS.

B. The Host Intrusion Detection System

According to the source of the data to examine, the Host Based Intrusion Detection System can be classified in two categories:

- The HIDS Based Application. The IDS of this type receive the data in application, for example, the logs files generated by the management software of the database, the server web or the firewalls. The vulnerability of this technique lies in the layer application.
- The HIDS Based Host. The IDS of this type receive the information of the activity of the supervised system. This information is sometimes in the form of audit traces of the operating system. It can also

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

include the logs system of other logs generated by the processes of the operating system and the contents of the object system not reflected in the standard audit of the operating system and the mechanisms of logging. These types of IDS can also use the results returned by another IDS of the Based Application type.

C. The Systems Detection Intrusion Hybrids

The NIDS-HIDS combination or the so called hybrid gathers the features of several different IDS. It allows, in only one single tool, to supervise the network and the terminals. The probes are placed in strategic points, and act like NIDS and/or HIDS according to their sites. All these probes carry up the alerts then to a machine which centralize them all, and aggregate the information of multiple origins.

III. INTRUSIONS PREVENTION SYSTEM

The intrusion prevention is an amalgam of security technologies. Its goal is to anticipate and to stop the attacks [2]. The intrusion prevention is applied by some recent IDS. Instead of analyzing the traffic logs, which lies in discovering the attacks after they took place, the intrusion prevention tries to warn against such attacks. While the systems of intrusion detection try to give the alert, the intrusion prevention systems block the traffic rated dangerous.

Over many years, the philosophy of the intrusions detection on the network amounted to detect as many as possible of attacks and possible intrusions and to consign them so that others take the necessary measures. On the contrary, the systems of prevention of the intrusions on the network have been developed in a new philosophy_ "taking the necessary measures to counter attacks or detectable intrusions with precision ".

In general terms, the IPS are always online on the network to supervise the traffic and intervene actively by limiting or deleting the traffic judged hostile by interrupting the suspected sessions or by taking other reaction measures to an attack or an intrusion. The IPS functions symmetrically to the IDS; in addition to that, they analyze the connection contexts, automatize the logs analysis and suspend the suspected connections. Contrary to the classic IDS, the signature is not used to detect the attacks. Before taking action, The IDS must make a decision about an action in an appropriate time. If the action is in conformity with the rules, the permission to execute it will be granted and the action will be executed. But if the action is illegal an alarm is issued. In most cases, the other detectors of the network will be informed with the goal to stop the other computers from opening or executing specific files.

Unlike the other prevention techniques, the IPS is a relatively new technique. It is based on the principle of integrating the heterogeneous technologies: firebreak, VPN, IDS, anti-virus, anti-Spam, etc.

The IPS are often considered as IDS of second generation; that is to say, the IPS replace the IDS gradually. In fact, the IPS are meant to make up for the limitations of the IDS concerning attacks response. Whereas the IDS cannot block an intrusion if it is not via the use of active responses, the IPS are able to block an intrusion in the appropriate time. Indeed, the positioning of the cut, be it in a firewall or in a proxy, is the only means which allows to analyze the input and output data and to destroy the intrusive packets dynamically before they arrive to their destination. Moreover, the IPS enable to compensate the IDS inability to manage the high debits because of a software architecture.

The IPS allow the following functionalities [8]:

- Supervising the behaviour of the application
- Creating rules for the application
- Issuing alerts in case of violations
- Correlating different sensors to guarantee a better protection against the attacks.
- Understanding of the IP networks
- Having mastery over the network probes and the logs analysis
- Defending the vital functions of the network
- Carrying out an analysis with high velocity.

A. The Network Intrusion Prevention System

When the attack is detected, the system reacts to modify the environment of the attacked system. This modification can be in the form blocking some fluxes and some ports or in the form of insulating some network systems. Directly affected system traffic is the sensitive point of this kind of prevention device especially when the false is positive. Therefore, the mistakes must be few because they have a direct impact on the availability of the systems. When dangerous traffic is detected, the IPS blocks this traffic like a firewall. Nevertheless, the same traffic, which takes place in a non dangerous configuration, won't be blocked. An IPS can be seen as identical to an intelligent firewall with dynamic rules [7].

B. The Host Intrusion Prevention System

Nowadays, the attacks evolve quickly and are targeted. Also, it is necessary to have a protection capable to stop the malwares before the publication of an update of the specific detection. An intrusions prevention system based on the Host Intrusion Prevention System or HIPS is destined to stop the malwares before an update of the specific detection is taken by supervising the code behaviour. The majority of the HIPS solutions supervises the code at the time of its execution and intervenes if the code is considered suspected or malevolent [7].

IV. FEATURES TO EVALUATE AND TO COMPARE FOR THE IDS/IPS SYSTEMS

The expression" system of detection and prevention of the intrusions" is used to describe multiple technologies

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

and solutions of security. This paper focuses on the systems of prevention of the intrusions capable to take immediate measures to tackle the attacks and intrusions without manual intervention. The tools of the intrusions detection and prevention systems display the following features:

a. Online machine capable to reliably and accurately detect the attacks and to block them with precision
b. High online velocity without any effect on the performance or the availability of the network
c. Efficient integration within the environment of the security management
d. Easy and quick adaptation with and anticipation of the unknown intrusions
e. Accurate and precise intervention
f. Good citizenship on the network
g. Efficient security-based management

An IDS/IPS system must include flexible and transparent methods to update its data-base with regard to the new signatures of attack. Besides, the IDS/IPS systems must have methods capable to react to new attacks without updates of signature.

The inverse exclusion, where all requests, except of those legitimate for a definite destination, are deleted, the validation of protocol, in which the methods of illegitimate requests are deleted, or the independent blockage of the attack, where the attackers are identified and the whole traffic that comes is deleted, whether the attacks are known or not.

V. THE FEATURES OF CLASSIFICATION OF THE IDS AND THE IPS.

There are a lot of products whose complexity of implementation and degree of integration are varied. The tools strictly based on behavioural models affect the velocity. But they are more and more integrated in IDS / IPS initially based on a library of signatures, thanks to their complementarily. The tools systems are worst facing to the tools networks. The invention of the hybrid tools that brings a less partial security in the protection of the system of information can solve this dilemma.

The first criterion of classification of the IDS/IPS is the method of analysis. It consists in two approaches.

- The approach by script: this approach consists in searching for in the activity of the element supervised the prints (or signatures) of known attacks. This type of IDS/IPS is merely reactive; it can only detect the attacks of which it possesses the signature. Therefore, it requires frequent updates. Besides, the efficiency of this detection system depends strongly on the precision of its signature basis. This is why these systems are vulnerable for the pirates who use some techniques “escape" that consists in making up the used attacks. These techniques have the trend to vary the signatures of the attacks that are not recognized anymore by the IDS/IPS
- The behavioural approach: it consists in detecting some anomalies. The implementation always consists of a phase of training during which the IDS/IPS is going *to discover* the *normal* functioning of the supervised elements. They are able, thus, to signal the divergences in relation to the working of the reference. The behavioural models can be elaborated from statistical analyses. They present the advantage to detect new types of attacks. However, frequent adjustments are necessary in order to evolve the reference model so that it reflects the normal activity of the users and reduce the number of false alerts generated.

Each of these two approaches can drive to *false positives* or to *false negatives*.

The intrusion detection and prevention systems become indispensable at the time of the setting up of an operational security infrastructure. Therefore, they always integrate in a context and in an architecture imposing various constraints.

The following criteria will be adopted in the classification of the IPS/IDS:

- *Reliability*: The generated alerts must be justified and no intrusion to escape
- *Reactivity*: An IDS/IPS must be capable to detect and to prevent the new types of attacks as quickly as possible. Thus, it must constantly self-update. Capacities of automatic update are so indispensable
- *Facility of implementation and adaptability*: An IDS/IPS must be easy to function and especially to adapt to the context in which it must operate. It is useless to have an IDS/IPS giving out some alerts in less than 10 seconds if the resources necessary to a reaction are not available to act in the same constraints of time
- *Performance*: the setting up of an IDS/IPS must not affect the performance of the supervised systems. Besides, it is necessary to have the certainty that the IDS/IPS has the capacity to treat all the information in its disposition because in the reverse case it becomes trivial to conceal the attacks while increasing the quantity of information.

These criteria must be taken into consideration while classifying an IDS/IPS, as well:

- The sources of the data to analyze, *network*, *system* or *application*
- The behaviour of the product after intrusion ,*passive* or *active*
- The frequency of use, *periodic* or *continuous*
- The operating system in which operate the tools, *Linux*, *Windows*, *etc*.
- The source of the tools, *open* or *private*

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

VI. THE TOOL IDS / IPS

In order to ensure an invulnerable security of data, various tools are available. They are mainly used altogether in order to secure the system as a whole. To avoid all sorts of inconveniences of the NIDS, NIPS, HIDS or HIPS it is very important to combine these different systems. The lack of information at the host level of the NIDS and NIPS in addition to the cost of installation-administration of the HIDS can be overcome through a good cohabitation of these systems on the network. There is no perfectly complete system. The optimum security is achieved as a result of the combination of several systems.

Moreover, most of these solutions are developed by the leading companies of securities. These solutions are complete and can be easily put in work in a network, which is also true for the updates. The modular format used by these allows them to have several agents for a centralized interface. However, these solutions are particularly very expensive.

Most of the existing solutions concerning intrusion detection are related to the setting up of NIDS in association with some HIDS and other software types of management.

The table below shows a study of the most used solutions of detection and prevention in the domains of commerce and open sources.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

Tools	CA eTRUST Intrusion Detection 3.0	Juniper IDP	McAfee Intrushield série I	McAfee Entercept 5.0	Snort 2.1.3	SonicWALL IPS service
Analysis of real-time traffic	Yes	Yes	Yes	Yes	Yes	Yes
Detection of viruses / worms / Trojans	Yes	Yes	Yes	Yes	Yes	Yes
Detecting external attacks	Yes	Yes	Yes	Yes	Yes	Yes
Detection of internal attacks	Yes	Yes	Yes	Yes	Yes	Yes
Ability to block attacks	Yes	Yes	Yes	Yes	Yes	Yes
Detection of external probes	Yes	Yes	Yes	Yes	Yes	Yes
Detection of internal Probes	Yes	Yes	Yes	Yes	Yes	Yes
Probes Ability	Yes	Yes	Yes	Yes	Yes	Yes
Definitions of blocking	Yes	Signatures with state data, protocol anomaly detection, backdoors, abnormal traffic, protection of layer 2, Syn Flood, Profiling enterprise security	Updates, block lists and user-defined customizable rules	Updates, block lists and user-defined customizable rules	Update, third-party integration, user-customizable	Updates
Real-time alert	E-mail, pager, application performance, SNMP, console	E-mail, syslog, SNMP, log file, external SMS	Console, email, pager, SMS email	Console, email, pager, SNMP, generation of process	Log files, email, console, third-party applications	Log files, email, syslog, SGMS

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

Getting logs data packets	Workspace, ODBC database	Syslog, internal database	Oracle, MySQL	Microsoft SQL Server	SS	SS
Search for content	Yes	Yes	SS	SS	Yes	Yes
Content Filtering	Yes	Yes	SS	SS	Yes	Yes
Filtering methods	URL database	Set by the administrator	SS	SS	Set by the administrator	Blacklist, third, set by the administrator
Reporting tools	Yes	Yes	Yes	Yes	SS (sold separately)	SS (sold separately)
Compatible operating system	Win 2000, Win 2000/2003/XP for the engine remotely	Windows, Linux, Solaris	Windows	Windows, Solaris, HP/UX	Linux, Windows	All IP environment

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

VII. CONCLUSION

With the multiplication of the networks of enterprise and the importance of Internet for the consumer, the enterprises try to make more and more present and visible on Internet. This presence on Internet, that it is through Internet sites, of the on line sale or even the mail often gets used to the detriment of the security of the networks of the enterprise and the data of the enterprise. As we saw it, many systems permit to reinforce the security on the networks of enterprise. That it is the firewalls, which filters the entry of the networks, the NIDS, that control through their probes, of the precise points of the networks, the HIDS, that supervise the intrusions directly at the host, or even the NIPS that have the capacity not to react at the time of the detection of activities dangerous, no system constitute the miracle remedy to the threatens computer attack.

Because of the inherent limits to each of these systems or techniques known of bypassing of these systems, the best protection was constituted of a combination of all these systems.

The versions of these protective systems are proposed commercially by different societies or organizations, under shape owner or free. According to the size of the enterprises and the means of these, there are some private solutions very easy of installation and configuration but unfortunately very expensive, some free and little expensive solutions also exist but unfortunately more difficult to install and to configure. The definition of the needs is therefore an indispensable preliminary stage before setting up these types of systems.

Besides, these systems can only act in the setting of a complement to a global security politics in all the enterprise, and constitute a small part of the security infrastructure.

The formation of the users but also of the administrators is also an indispensable point to this politics.

In order to improve the capacities of control and protections of these systems, the research are always in progress. These researches try to optimize the present systems or to find new solutions of detection, filtering or reaction after alert.

Some firewall or firewall integrating the IDS or the IPS appear, even for the general public level. The democratization of these types of systems permits, gradually, to bring a beginning of security, that was not often considered important by the decision-makers in the past. In a general manner, the efficiency of a system of intrusion detection depends on its "configurability" (possibility to define and to add new specifications of attack), of its hardiness (resistance to the failings) and of the quantity of false positives (false alerts) and of false negatives (non detected attacks) that it generates. The paragraphs have at a time for objectives to illustrate the complexity of intrusion detection and to explain the limits of the present IDS. A struggle between techniques of intrusion and IDS began, the IDS having for consequence a bigger technicality of the attacks on IP, and the present attacks imposing to the IDS to be more complete and more powerful [8]. The IDS/IPS bring an incontestable advantage to the networks in which they are placed. However, their limits don't permit to guarantee a security to 100%, impossible to get. The future of these tools will permit to fill these hiatuses by avoiding the "false positives" (for the IDS) and refining the restrictions of access (for the IPS) "[5].

This study has proved that both the intrusion detection systems and the intrusion prevention systems still need to be improved to ensure an unfailing security for a network. They are not reliable enough (especially in regard to false positives and false negatives) and they are difficult to administer. Yet, it is obvious that these systems are now essential for companies to ensure their security. To assure an effective computerized security, it is strongly recommended to combine several types of detection system. The IPS, which attempt to compensate in part for these problems, are not yet effective enough for use in a production context. They are currently mainly used in test environments in order to evaluate their reliability. They also lack a normalized operating principle like for the IDS. However, these technologies require to be developed in the coming years due to the increasing security needs of businesses and changes in technology that allows more efficient operation detection systems and intrusion prevention. We are working on the implementation of a screening tool of attack and the characterization of test data. We also focus on the collection of exploits and attacks to classify and identify. Further work is under way and many ways remain to be explored. Then it would be interesting to conduct assessments of existing IDS and IPS following the approaches we have proposed and tools developed in this work.

.

REFERENCES

[1] Crying wolf: False alarms hide Newman attacks, Snyder & Thayer Network World, 24/06/02, http://www.nwfusion.com/techinsider/2002/0624security1.html

[2] F. Cikala, R. Lataix, S. Marmeche", The IDS/IPS. Intrusion Detection/Prevention Systems ", Presentation, 2005.

[3] Hervé Debar and Jouni Viinikka, "Intrusion Detection,: Introduction to Intrusion Detection Security and Information Management", Foundations of Security Analysis and Design III, Reading Notes in to Compute Science, Volume 3655, 2005. pp. 207-236.

[4] Hervé Debar, Marc Dacier and Andreas Wespi, "IN Revised Taxonomy heart Intrusion Detection Systems", Annals of the Telecommunications, Flight. 55, Number,: 7-8, pp. 361-378, 2000.

[5] Herve Schauer Consultants", The detection of intrusion…", Presentation: excerpt of the course TCP/IP security of the Cabinet HSC, March 2000.

[6] ISS Internet Risk Impact Summary - June 2002.

[7] Janne Anttila", Intrusion Detection in Critical Ebusiness Environment ", Presentation, 2004.

[8] D K. Müller", IDS - Systems of intrusion Detection, Left II ", July 2003, http://www.linuxfocus.org/Francais/July2003/article294.shtml

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

Network Intrusion Detection Types and Computation

Purvag Patel, Chet Langin, Feng Yu, and Shahram Rahimi
Southern Illinois University Carbondale, Carbondale, IL, USA

***Abstract*—Our research created a network Intrusion Detection Math (ID Math) consisting of two components: (1) a way of specifying intrusion detection types in a manner which is more suitable for an analytical environment; and (2) a computational model which describes methodology for preparing intrusion detection data stepwise from network packets to data structures in a way which is appropriate for sophisticated analytical methods such as statistics, data mining, and computational intelligence. We used ID Math in a production Self-Organizing Map (SOM) intrusion detection system named ANNaBell as well as in the SOM+ Diagnostic System which we developed.**

***Index Terms*—Computational intelligence, Data Mining, ID Math, Intrusion Detection Types, Log Analysis**

I. Introduction

Every hacker in the world is one's neighbor on the Internet, which results in attack defense and detection being pervasive both at home and work. Although hundreds of papers have been written on a large variety of methods of intrusion detection—from log analysis, to packet analysis, statistics, data mining, and sophisticated computational intelligence methods—and even though similar data structures are used by the various types of intrusion analysis, apparently little has been published on a methodical mathematical description of how data is manipulated and perceived in network intrusion detection from binary network packets to more manageable data structures such as vectors and matrices.

We developed a comprehensive methodology of information security Intrusion Detection Math (ID Math) which overhauls concepts of intrusion detection including a new model of intrusion detection types and a computational model created in order to lay a foundation for data analysis. Our intrusion detection types are necessary, complete, and mutually exclusive. They facilitate *apples-to-apples* and *oranges-to-oranges* comparisons of intrusion detection methods and provide the ability to focus on different kinds of intrusion detection research. Our computational model converts intrusion detection data from packet analysis step-by-step to sophisticated computational intelligent methods. These concepts of ID Math were implemented in a production Self-Organizing Map (SOM) intrusion detection system named ANNaBell and were introduced in publication as part of the SOM+ Diagnostic System in [1].

Section II describes background and literature. We describe the new types of local network intrusion detection in section III, and we propose the network intrusion detection computation model in section IV. The conclusion is in section V.

II. Background and Literature

Intrusion detection is the process of identifying and responding to malicious activity targeted at computing and networking sources [2]. Over the years, types of intrusion detection have been labeled in various linguistic terms, with often vague or overlapping meanings. Not all researchers have used the same labels with the same meanings. To demonstrate the need for consistent labeling of intrusion types, previous types of intrusion detection are listed below in order to show the variety of types of labeling that have been used in the past.

Denning [3] in 1986 referred to intrusion detection methods which included profiles, anomalies, and rules. Her profiling included metrics and statistical models. She referred to misuse in terms of insiders who misused privileges.

Young in 1987 [4] defined two types of monitors: appearance monitors and behavior monitors, the first performing static analysis of systems to detect anomalies and the second examining behavior.

Lunt [5] in 1988 referred to the misuse of insiders; the finding of abnormal behavior by determining departures from historically established norms of behavior; a priori rules; and using expert system technology to codify rules obtained from system security officers. A year later, in 1989, Lunt mentioned knowledge-based, statistical, and rule-based intrusion detection. In 1993, she referred to model-based reasoning [6].

Vaccaro and Liepins [7] in 1989 stated that misuse manifests itself as anomalous behavior. Hellman, Liepins, and Richards [8] in 1992 stated that computer use is either normal or misuse. Denault, et al, [9] in 1994 referred to detection-by-appearance and detection-by-behavior. Forrest, et al, [10] in 1994 said there were three types: activity monitors, signature scanners, and file authentication programs.

Intrusion detection types began converging on two main types in 1994: misuse and anomaly. Crosbie and Spafford [11] defined misuse detection as watching for certain actions being performed on certain objects. They defined anomaly detection as deviations from normal system usage patterns. Kumar and Spafford [12] also referred to anomaly and misuse detection in 1994. Many other researchers, too numerous to mention them all, have also referred to misuse and anomaly as the two main types of intrusion detection, from 1994 up to the present time.

However, other types of intrusion detection continue to be mentioned. Ilgun, Kemmerer, and Porras [13] in 1995 referred to four types: Threshold, anomaly, rule-based, and model-based. Esmaili, Safavi-Naini, and Pieprzyk [14] in 1996 said the two main methods are statistical and rule-based expert systems.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

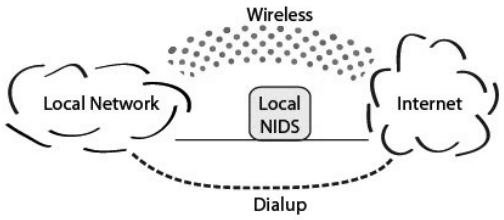


Fig. 1. A Local Landline NIDS

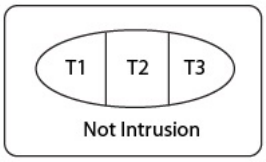


Fig. 2. Types of Intrusions for LLNIDS

Debar, Dacier, and Wespi, [15] in 1999 referred to two complementary trends: (1) The search for evidence based on knowledge; and, (2) the search for deviations from a model of unusual behavior based on observations of a system during a known normal state. The first they referred to as misuse detection, detection by appearance, or knowledge-based. The second they referred to as anomaly detection or detection by behavior. Bace [16] in 2000 described misuse detection as looking for something bad and anomaly detection as looking for something rare or unusual. Marin-Blazquez and Perez [17] in 2008 said that there are three main approaches: signature, anomaly, and misuse detection.

While descriptive, these various labels over time are inconsistent and do not favor an analytical discussion of network intrusion detection. Not all of them are necessary, they are not mutually exclusive, and as individual groups they have not been demonstrated as being complete. Rather than arbitrate which of these labels should be used and how they should be defined, new labels have been created to describe types of local network intrusion detection in a manner which favors an analytical environment.

III. LLNIDS Types of Intrusion Detection

The new types are explained below, but first some terminology needs to be stated in order to later describe the types. An Intrusion Detection System (IDS) is software or an appliance that detects intrusions. A Network Intrusion Detection System (NIDS) is an appliance that detects an intrusion on a network. In this research, network means a landline network. Local network intrusion detection refers to the instant case of network intrusion detection.

Figure 1 illustrates the location of a Local Landline Network Intrusion Detection System (LLNIDS) as used in this research. The LLNDS in Figure 1 is represented by the rounded box in the center labelled “Local NIDS”. It is an IDS on a landline between a local network and the Internet. The point of view of this research is from inside the LLNIDS. Users on the local network may have other ways of accessing the Internet that bypass the LLNIDS, such as wireless and dialup. This research is restricted to the LLNIDS as described here.

Examples of detection which are not Local Landline Network Intrusion Detection (LLNID) include detection on the host computer, detection by someone else out on the Internet, or detection by someone out in the world, such as someone witnessing a perpetrator bragging in a bar. This research concerns LLNID and the new types described in this paper refer to LLNID. A network intrusion in this context means one or more transmissions across the network that involves an intrusion. A single Internet transmission is often called a packet. Therefore, using this terminology, the physical manifestation of an intrusion on a network is one or more packets, and intrusion detection is the detection of these packets that constitute intrusions. In this context, intrusion detection is similar to data mining. Intrusion detection research needs a model of types of intrusions and types of intrusion detection that benefits analysis of methods. This research focuses only on LLNID. These are the proposed types of intrusions for the special case of local landline network intrusion detection that facilitate intrusion detection research analysis in the LLNID context:

- *Type 1 Intrusion*: An intrusion which can be positively detected in one or more packets in transit on the local network in a given time period.
- *Type 2 Intrusion*: An intrusion for which one or more symptoms (only) can be detected in one or more packets in transit on the local network in a given time period.
- *Type 3 Intrusion*: An intrusion which cannot be detected in packets in transit on the network in a given time period.

These three types of intrusions are necessary for analytical research in order to indicate and compare kinds of intrusions. A positive intrusion is different than only a symptom of an intrusion because immediate action can be taken on the first whereas further analysis should be taken on the second. Both of these are different than intrusions which have been missed by an LLNIDS. To show that these three types are mutually exclusive and are complete for a given time period, consider all of the intrusions for a given time period, such as a 24-hour day. The intrusions which were positively identified by the LLNIDS are Type1 intrusions. Of the remaining intrusions, the ones for which the LLNIDS found symptoms are Type 2. Here the hypothesis is that the LLNIDS can only find an intrusion positively or only one or more symptoms are found. No other results can be returned by the LLNIDS. Therefore, the remaining intrusions are Type 3, which are intrusions not detected by the LLNIDS. No other types of intrusions in this context are possible.

Figure 2 is a diagram that illustrates the types of intrusions as described above. An intrusion is either Type 1, Type 2, Type 3, or it is not an intrusion.

Those were the types of intrusions. Next are the types of intrusion detection. There are three types of network intrusion detection that correspond to the three types of intrusions in the LLNID context:

- *Type 1 Network Intrusion Detection*: A Type 1 Intrusion is detected in a given time period.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

- *Type 2 Network Intrusion Detection*: One or more symptoms (only) of a Type 2 Intrusion are detected in a given time period.
- *Type 3 Network Intrusion Detection*: No intrusion is detected in a given time period.

Admittedly, Type 3 is not a detection but the lack of detection. It is included because these three types of detection correspond to the three types of intrusions and Type 3 Intrusion Detection facilitates analysis of intrusion detection methods. Examples of Type 3 Intrusion Detection are nothing was detected; no attempt was made at detection; an intrusion occurred but was not detected by the LLNIDS; and, no intrusion occurred. All of these have the same result: there was no detection of an intrusion by the LLNIDS.

Each of the three network intrusion detection types is necessary to describe all of the types of intrusion detection. A positive detection of an intrusion is different than just a symptom of an intrusion because a positive detection can be immediately acted upon while a symptom indicates that further analysis is needed. Both of these are different than intrusions that are missed by network intrusion detection. To show that these types are mutually exclusive and complete for a given time period, consider an LLNIDS looking at network packets for a given time period, say a 24-hour day. For all packets that the LLNIDS determines positively indicates an intrusion the LLNIDS has accomplished Type 1 intrusion detection. Of the remaining packets, for each packet that the LLNIDS determines is a symptom of an intrusion the LLNIDS has accomplished Type 2 intrusion detection. The remaining packets represent Type 3 intrusion detection. These three types of network intrusion detection are complete in this context because they cover all possibilities of intrusion detection. In common language, Type 1 is a certainty, Type 2 is a symptom, and Type 3 is an unknown.

Those were types of intrusion detection. Next are types of methods and alerts. LLNID methods can be defined in terms of the three intrusion types:

- *Type 1 NID Method/Alert*: A method that detects a Type 1 Intrusion and an alert that indicates a Type 1 Intrusion.
- *Type 2 NID Method/Alert*: A method that detects a symptom of a Type 2 Intrusion and an alert that indicates a symptom (only) of a Type 2 Intrusion.
- *Type 3 NID Method/Alert*: A method that does not exist, thus there is no alert.

These types of methods and alerts are necessary to differentiate that some methods are positively correct, other methods only indicate symptoms of intrusions, and some methods do not exist. They are mutually exclusive because a local method either positively indicates an intrusion (Type 1), it only detects a symptom of an intrusion (Type 2), or it does not exist (Type 3). They are complete because there are no other types of methods in this context.

Those were types of methods and alerts. Next are types of false positives. The term false positive generally has meant that an intrusion detection system has sent a false alarm. False positives are generally undesirable because the false positive rate of intrusion detection systems can be high and can use up a lot of seemingly unnecessary, and limited, resources. However, with these new types, the concept of a false positive is different for different intrusion detection types in the LLNIDS context.

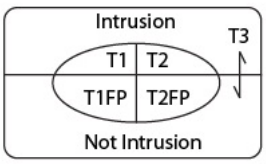


Fig. 3. Types of Intrusion Detection for LLNID

- *Type 1 False Positive*: A Type 1 Method produces an alarm in the absence of an intrusion.
- *Type 2 False Positive*: A Type 2 method produces an alarm in the absence of an intrusion.
- *Type 3 False Positive*: Does not exist because no alarm is produced.

A Type 1 False Positive indicates a problem with the Type 1 method which should be corrected. Type 2 False Positives are expected because Type 2 Methods do not positively detect intrusions, they only detect symptoms of intrusions. There is no Type 3 False Positive because no detections and alerts are produced for Type 3 Intrusion Detections. These types of false positive are necessary because they each indicate separate network intrusion detection issues. Type 1 is a network intrusion detection problem which needs to be corrected and Type 2 is expected. The two types of false positive are mutually exclusive and complete because only Type 1 Network Intrusion Detection can produce a Type 1 False Positive and only Type 2 Network Intrusion Detection can produce a Type 2 False Positive. No other types of false positives in this context are possible. Since Type 1 and Type 2 of local network intrusion detection methods are mutually exclusive, these are also mutually exclusive.

Figure 3 is a Venn diagram which illustrates types of intrusion detection in the LLNIDS context. The horizontal line separates intrusions at the top from non-intrusions at the bottom. A Type 1 detection is in the upper left of the circle if it is actually an intrusion or it is in the lower left of the circle if it is a false positive. A Type 2 detection is in the upper right of the circle if it is actually an intrusion or it is in the lower right of the circle if it is a false positive. Everything outside of the circle is Type 3 detection whether it is an intrusion or not.

This typing system allows illustration that empirically most intrusion detection is not Type 1 (positive detections), but Type 2 (symptoms of detections), and Type 3 (missed detections). This differentiation is essential in proceeding in a scientific way for improved intrusion detection.

Previously labeled types of intrusion detection do not fit neatly into these three new types. Misuse detection, for example, in some cases could indicate a definite intrusion and would then be Type 1, or it could indicate only symptoms of intrusions in other cases and would then be Type 2. The comparison of false positives of different methods of Misuse Detection is an invalid technique unless Type 1

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

methods are compared only with Type 1 methods and Type 2 methods are compared only with Type 2 methods. Anomaly detection, for example, would tend to be Type 2, but some anomalies could clearly indicate intrusions and would be Type 1. Type 1 and Type 2 methods of Anomaly Detection should be separated before making any comparisons. Likewise with intrusion detection labels based on activity, appearance, authentication analysis, behavior, knowledge, models, profiles, rules, signature, static analysis, statistics, and thresholds. These are still useful as descriptive terms, but they are not as useful in analyzing methods of determining whether or not an intrusion has occurred because they allow the comparisons of *apples* and *oranges* in numerous ways. The labels Type 1 and Type 2 give us more analytical information: either an intrusion has occurred or else only a symptom of an intrusion has occurred. Type 3 intrusions tell us that we should find out why an intrusion was not detected in the network traffic so that we can create new rules to find more intrusions in the future. Previously labeled types of intrusion detection do not give us as much analytical information as do types 1, 2, and 3.

Using this system, one can clearly state objectives of LLNID research in a new way which was previously only implied. The significance of given time period is apparent in the descriptive of these objectives because the objectives are stated in terms of progress from one time period to another time period. Here are specifics for LLNID research:

- *Type 3 NID Research*: Find ways of detecting intrusions that are currently not being detected, moving them up to type 2 or 1 intrusion detection.
- *Type 2 NID Research*: Improve Type 2 Intrusion Detection with the goal of moving it up to Type 1 Intrusion Detection.
- *Type 1 NID Research*: Improve Type 1 Intrusion Detection so that it is faster, uses fewer resources, and has fewer false positives.

Each of these types of research are necessary because finding new methods of intrusion detection is different than improving symptom detection which is different than making Type 1 Intrusion Detection more efficient. They are also complete because there are no other types of intrusion detection research in this context.

Table 1 summarizes the types discussed in this section. These are some ways of how researchers can use these types: research that compares false positive rates of Type 1 methods with false positive rates of Type 2 methods is not valid because Type 1 methods are not supposed to have false positives whereas Type 2 methods are expected to have false positives. Discounting Type 3 intrusion detection because of the amount of time taken may be irrelevant if otherwise the intrusion would not be found, at all. Proposing that intrusion prevention will replace intrusion detection is a false claim so long as types 2 and 3 intrusions continue to exist. Rather than disregarding Type 2 methods, research should attempt to fuse the results of Type 2 methods in order to move them up to Type 1.

TABLE I
SUMMARY OF LLNID TYPES

	Type 1	Type 2	Type 3
Intrusion	This can be positively detected by LLNIDS	A symptom of this can be detected by LLNIDS	This is not detected by LLNIDS
Intrusion Detection	This positively detects an intrusion	This detects one or more symptoms (only) of an intrusion	An intrusion is not detected
Method	How to positively detect an intrusion	How to positively detect a symptom of an intrusion	An intrusion is not detected
Alert	This positively signifies an intrusion	This signifies a symptom of an intrusion	This does not occur
False Positive	An alert positively signifies an intrusion, but there is no intrusion	An alert signifies a symptom of an intrusion, but there is no intrusion	An alert does not occur
Research	Improve Type 1 Intrusion Detection, such as by increasing the speed of detection, using less resources, and having fewer false positives	Improve Type 2 Intrusion Detection so that it becomes Type 1 Intrusion Detection	Detect Type 3 intrusions so that they become Type 2 or Type 1

IV. THE LLNIDS COMPUTATIONAL MODEL

A few number of researchers have described intrusion detection in limited mathematical ways, with [18][19], in the context of attack trees, and [20], in the context of game theory, being representative. Network Monitoring was formulated as a language recognition problem in [21].

We propose Local Landline Network Intrusion Detection System (LLNIDS) Computational Model that covers intrusion detection data from packet analysis to sophisticated computational intelligent methods. This ID Math computational model begins with a transmission of digital network traffic and proceeds stepwise to higher concepts. The terminology for the input data changes depending upon the level of the concept. The lowest level concept in this research is the network transmission, which is a series of bits called a frame or a packet. Frame refers to a type of protocol, such as Media Access Control (MAC), which is used between two neighboring devices, where the series of bits are framed by a header at the start and a particular sequence of bits at the end. Packet refers to many types of protocols, such as Internet Message Control Protocol (ICMP), User Datagram Protocol (UDP), and Transmission Control Protocol (TCP). A packet is used for hops between numerous devices, such as Internet traffic. The length of the series of bits in a packet is often indicated at certain locations in the headers of the packets. A frame passes a packet between two neighboring devices, where another frame passes the same packet between the next two devices, and subsequent frames keep passing the packet forward until the journey of the packet is concluded. Since frames and packets are variable lengths, they are represented by a set of objects which represent the various elements of information inside the frame or packet.

A Transmission (T) consists of a set of objects (o) representing elements of information in that transmission.

$$T = \{o_1, o_2, o_3, \ldots, o_{n_T}\} \quad (1)$$

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

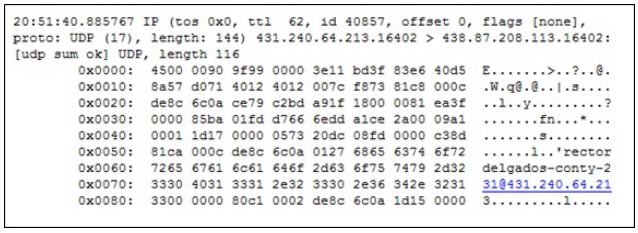


Fig. 4. A Sample Packet

where $n_T \in \mathcal{N}$. Examples of objects in a transmission are the source MAC address, source IP address, source port, destination MAC address, destination IP address, destination port, the apparent direction of the traffic, protocols used, flags set, sequence numbers, checksums, type of service, time to live, fragmentation information, and the content being sent.

Figure 4 is a sample packet as displayed by tcpdump [22]. Header information extracted from the packet is displayed across the top. The leftmost column is the byte count in hexadecimal. The packet itself is displayed in hexadecimal in columns in the middle. Character representations of the hexadecimal code, when possible, are shown on the right. The packet is a transmission set, T, with variable length objects as elements. Example object elements for this set are the protocol, UDP, and the destination port, 16402, both of which have been extracted from the packet code.

If an intrusion occurs on a local landline, it occurs in one or more T, so LLNID means inspecting T's for intrusions. Not all of the available data in T has equal relevance to intrusion detection and the reduction of the amount of data is desirable in order to reduce the resources needed for analysis. This process has been called feature deduction [23], feature reduction [23], feature ranking [24], or feature selection [23]. The first feature selection must be done manually by a knowledge engineer, after that the features can be ranked and/or reduced computationally. Soft Computing methods often use data structures of n-tuple formats, such as one-dimensional arrays, sets, vectors, and/or points in space. Since sets can be used as a basis to describe these data structures, the next step in the computational model is to convert features of T into higher levels of sets which can be further manipulated for data analysis. The next set to be considered is an Event (E) which consists of a set of elements (e) obtained from the objects of T, and which changes the concept level from a transmission of objects to a set of elements:

$$E = \{e_1, e_2, e_3, \ldots, e_{n_E}\} \tag{2}$$

where $n_E \in \mathcal{N}$ and the following condition is also met:

$$\forall e_i \in E, 1 \leq i \leq n_E, e_i \in T \tag{3}$$

How to construct e_i from the objects of T is feature selection–elements should be selected which can detect intrusions. An example of possible elements for an event is the source IP address, the destination IP address, the source and destination ports, the protocol, and the size of a packet crossing the network.

TABLE II
A SAMPLE EVENT

UDP	231.240.64.213	238.87.208.113	16402

TABLE III
SAMPLE META-DATA

20100916	00:14:54	FW

Table 2 shows a sample event with the following elements: The protocol is UDP, the source IP address is 231.240.64.213, the destination IP address is 238.87.208.113, and the destination port is 16402. These elements were object elements in the sample transmission set shown above. The process of pulling data objects from a packet and saving them as Event elements is called parsing the data.

The next step is to add Meta-data (M), if appropriate, about the event consisting of meta-data elements (m):

$$M = \{m_1, m_2, m_3, \ldots, m_{n_M}\} \tag{4}$$

where $n_M \in \mathcal{N}$. Meta-data is data about data. In this context, it means data about the transmission that is not inside the transmission, itself. Examples of meta-data are the time when a packet crossed the network, the device which detected the packet, the alert level from the device, the direction the packet was travelling, and the reason the packet was detected. The concept level has changed from a set of elements to a set of meta-data about the set of elements.

Table 3 shows sample meta-data for an event. The meta-data in this table is the date, 20100916, and the time, 00:14:54, at which an appliance detected the transmission, and a label for the appliance that detected the packet, FW.

A Record (R) of the event includes both the event, itself, plus the meta-data:

$$R = M \cup E \tag{5}$$

An example of a record is an entry in a normalized firewall log. The concept level has changed from a set of meta-data to a set that includes both the elements and meta-data about those elements. In practice, the meta-data typically occurs in R before the elements to which the meta-data refers.

Table 4 is a sample record, which consists of meta-data and elements from the previous examples for M and E. Before proceeding to the next step, the attributes of R for a given analysis should be in a fixed order because they can later become coordinates in a location vector. Processing the data into fixed orders of attributes is called normalizing the data.

A Log (L) of records is a partially ordered set:

$$L = \{R_i\}_{i \in \mathcal{N}} \tag{6}$$

An example of a log is a file containing normalized firewall log entries. An infinite-like log could be live streaming data.

TABLE IV
A SAMPLE RECORD

20100916	00:14:54	FW	UDP	231.240.64.213	238.87.208.113	16402

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

TABLE V
A SAMPLE LOG

20100916	00:14:54	FW	UDP	231.240.64.213	238.87.208.113	16402
20100916	00:14:56	FW	TCP	216.162.156.85	198.18.147.222	40833
20100916	11:14:57	FW	ICMP	90.29.214.20	198.18.147.221	41170

Table 5 shows a sample log. It is like the sample record, above, except there are three entries instead of just one entry. The concept level has changed from a set of meta-data and elements to a collection of sets of meta-data and elements. L can be considered to be a set of vectors; L can also be considered to be a matrix. If L is a text file, each line of the file is one location vector and the entire file is a matrix, changing the concept level to a matrix.

If the features have been selected successfully, an intrusion, or one or more symptoms of it, should be able to be detectable in L. Therefore, LLNIDS intrusions and intrusion detection can be defined in terms of R and L. Let $\mathscr{R}$ be the universal set of R and let I_1 represent a set of $\mathscr{R}$ that describe a Type 1 Intrusion. Then I_1 is the set:

$$I_1 = \{R | R \in \mathscr{R}, R \; involves \; a \; Type \; 1 \; Intrusion \; \} \quad (7)$$

Formula 7 formulates a Type 1 Intrusion. Examples of Type 1 intrusions are a Ping of Death and a get request to a known malicious web site. These intrusions can potentially be prevented. I_1 has the same attributes as L in that it can be considered to be a set of location vectors or it can be considered to be a matrix. As matrices, the number of columns in I_1 and L for an analysis must be the same, but the number of rows in I_1 and L can be different. For reference below, let $\mathscr{I}_1$ be the universal set of all Type 1 intrusions. The concept level for $\mathscr{I}_1$ has changed from a matrix to a set of matrices. That was about intrusions. Now here is the function for Type 1 Intrusion Detection, I_1^D:

$$I_1^D(L) = \begin{cases} \text{TRUE}, & \exists I_1 \in \mathscr{I}_1 : I_1 \subseteq L \\ \text{FALSE}, & otherwise \end{cases} \quad (8)$$

Formula 8 is the function for Type 1 Intrusion Detection, which returns True if an intrusion has been detected, otherwise it returns False. Next is Type 2 intrusions and intrusion detection. In most cases, one or more events occur which makes the security technician suspicious that an intrusion has occurred, but more investigation is necessary in order to reach a conclusion. This scenario, which is Type 2 Intrusion Detection, is similar to a patient going to a physician, who looks for symptoms and then makes a decision about whether or not the patient has a medical problem. The security technician also looks for symptoms and then makes a decision about whether or not an intrusion has occurred. Let $\mathscr{R}$ be the universal set of R and let I_2 represent a set of R that describes one or more symptoms of a Type 2 Intrusion. Then I_2 is the set:

$$I_2 = \{R | R \in \mathscr{R}, R \; involves \; a \; Type \; 2 \; Intrusion \; \} \quad (9)$$

Formula 9 formulates a Type 2 Intrusion. Let $\mathscr{R}_2$ be the universal set of all Type 2 intrusions. Now here is a formula for Type 2 Intrusion Detection, I_2^D:

$$I_2^D(L) = \begin{cases} \text{TRUE}, & \exists I_2 \in \mathscr{I}_2 : I_2 \subseteq L \\ \text{FALSE}, & otherwise \end{cases} \quad (10)$$

The $I_2^D(L)$ function returns True if a symptom of an intrusion has been detected; otherwise it returns False. Possible examples of Type 2 intrusions are the following: The set of records consisting of a single local source IP address and numerous unique destination addresses all with a destination port of 445; the set of records consisting of a local IP address sending numerous e-mails during non-working hours; and, the set of records consisting of high volumes of UDP traffic on high destination ports to a single local IP address matching criteria set by a Self-Organizing Map. Like a cough does not necessarily indicate a cold, the detection of an intrusion symptom does not always indicate an intrusion.

That was Type 2 intrusions and intrusion detection. Next is Type 3 intrusions, which are not detected in a given time period. Let $\mathscr{R}$ be the universal set of R and let I_3 represent a set of R that describes a Type 3 Intrusion. Then I_3 is the set:

$$I_3 = \{R | R \in \mathscr{R}, R \; involves \; a \; Type \; 3 \; Intrusion \; \} \quad (11)$$

As a summary, compare these three types of intrusion detection in a medical context to typhoid fever, which is spread by infected feces. Type 1 intrusion detection (prevention) is to wash one's hands after using the toilet; Type 2 intrusion detection is to recognize the symptoms, such as fever, stomach ache, and diarrhea; Type 3 detection is represented by Typhoid Mary, who had no readily recognizable symptoms.

The next step involves changing the data formats from R and L into forms which can be directly manipulated by analysis software. (Packet analysis can already occur directly on T.) This involves converting records into vectors and logs into matrices. This conversion is straightforward with a Detailed Input Data Vector, V_D, which starts as a set and is then used later as a location vector:

$$V_D \subseteq R \quad (12)$$

More feature reduction can occur at this step. If the order of each element in the set is fixed, i.e., if the order of the attributes of the set are fixed, then the set can become a location vector. An example of V_D as a set is {1280093999, 10.3.4.10, 10.3.4.12, 445, TCP} which could indicate a time stamp in seconds, a source IP address, a destination IP address, a destination port, and a protocol. Converting IP addresses to numerical formats, and assigning a numerical label to TCP, the same example of V_D as a location vector could be (1280093999, 167969802, 167969804, 445, 6).

Aggregate elements are also possible for a given time period, such as aggregate data for each local IP address for a day. Examples of such aggregate elements are the total number of R for the local IP address, the count of unique source IP addresses communicating with the local IP address, and the percentage of TCP network traffic for the local IP address. Many other types of aggregate elements are possible. These aggregate elements can be converted to an Aggregate Input Data Vector, V_A, with f being an aggregation function:

$$V_A = \{f_1(L), f_2(L), f_3(L), \ldots, f_{n_V}(L)\} \quad (13)$$

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

where $n_V \in \mathcal{N}$. Again, the order of the attributes of the set are fixed so that the set can become a location vector. An example of V_A as a set is {20100725, 428, 10.3.4.10, 48, 0.89} which could indicate that on 7/25/2010 428 unique source IP addresses attempted to contact destination IP address 10.3.4.10 on 48 unique destination ports with the TCP protocol being used 89 percent of the time. The date and IP address become a label for the location vector when the location vector is created. From the same example above, the location vector for IP address 10.3.4.10 on 7/25/2010 is (428, 48, 0.89).

Both of these types of sets/vectors can be generalized as a General Input Data Vector, V:

$$V = V_D \; or \; V = V_A \tag{14}$$

The next concept level is to generalize V so that it can be used as input to a wide variety of Soft Computer and other methods. The generalized elements of V are be represented by e. V is an n-tuple of real numbers which can be perceived, depending upon how it is intended as being used, as being a set, a location vector, or a matrix:

$$Set : V = \{e_1, e_2, e_3, \ldots, e_{n_V}\} \tag{15}$$

$$Vector : V = (e_1, e_2, e_3, \ldots, e_{n_V}) \tag{16}$$

$$Matrix : V = [e_1 \; e_2 \; e_3 \; \ldots \; e_{n_V}] \tag{17}$$

where $n_V \in \mathcal{N}$. For example, if the elements of V are an n-tuple of the real numbers 0.6, 0.5, 0.4, 0.3, 0.2, and 0.1, then V can be perceived as being a set, a vector or a matrix:

$$Set : V = \{0.6, 0.5, 0.4, 0.3, 0.2, 0.1\} \tag{18}$$

$$Vector : V = (0.6, 0.5, 0.4, 0.3, 0.2, 0.1) \tag{19}$$

$$Matrix : V = [0.6 \; 0.5 \; 0.4 \; 0.3 \; 0.2 \; 0.1] \tag{20}$$

An Input Data Matrix, D, is a collection of similar types of V. Here D is represented as a set of V:

$$D = \{V_1, V_2, V_3, \ldots, V_{n_D}\} \tag{21}$$

where $n_D \in \mathcal{N}$. D is on the same concept level as L. Both D and L can be considered to be sets of location vectors or a matrix. Here is how D can be represented as a matrix:

$$D = \begin{bmatrix} V_{1,1} & \cdots & V_{1,n_V} \\ \vdots & \ddots & \vdots \\ V_{n_D,1} & \cdots & V_{n_D,n_V} \end{bmatrix} \tag{22}$$

where $n_D \in \mathcal{N}$ and $n_V \in \mathcal{N}$.

For example, given these three location vectors, each represented as a matrix,

$$V_1 = [0.6 \; 0.5 \; 0.4 \; 0.3 \; 0.2 \; 0.1] \tag{23}$$

$$V_2 = [0.1 \; 0.2 \; 0.3 \; 0.4 \; 0.5 \; 0.6] \tag{24}$$

$$V_3 = [0.9 \; 0.8 \; 0.7 \; 0.6 \; 0.5 \; 0.4] \tag{25}$$

D would be represented this way as a matrix:

$$D = \begin{bmatrix} 0.6 & 0.5 & 0.4 & 0.3 & 0.2 & 0.1 \\ 0.1 & 0.2 & 0.3 & 0.4 & 0.5 & 0.6 \\ 0.9 & 0.8 & 0.7 & 0.6 & 0.5 & 0.4 \end{bmatrix} \tag{26}$$

D_D can refer to an Input Data Matrix consisting of V_D and D_A can refer to an Input Data Matrix consisting of V_A. D can also be one of these three types:

1) D_{Train} refers to a data set which is used to train the software intelligence
2) D_{Test} refers to a data set which is used to test the software intelligence
3) D_{Real} refers to feral data.

D can be used in virtually an infinite variety of analysis methods, from spreadsheet methods to statistics and data mining, to machine learning methods. For example, D_{Train} can be used by clustering software which, after testing, would then classify D_{Real} for intrusion detection.

The ID Math Method more accurately defines information security concepts and scientifically ties components of information security together with structured and uniform data structures. The LLNIDS can be extended to describe existing and potential methodologies of analysis methods including statistics, data mining, AIS, NeuroFuzzy, Swarm Intelligence, and SOM, as well as Bayes Theory, Decision Trees, Dempster-Shafer Theory, Evolutionary Computing, Hidden Markov Models, and many other types of analysis.

V. Conclusion

This paper provided a new way of looking at network intrusion detection research including intrusion detection types that are necessary, complete, and mutually exclusive to aid in the fair comparison of intrusion detection methods and to aid in focusing research in this area. This paper also provided a methodical description of intrusion detection data and how this data is manipulated and perceived from packet analysis to sophisticated computational intelligence methods. This new ID Math provides a methodological archetype from which to move forth. Future work in intrusion detection research should leverage these intrusion detection types and this computational model for better descriptions of the problem sets and for presenting solutions to intrusion detection.

References

[1] Langin, C. L. A SOM+ Diagnostic System for Network Intrusion Detection. Ph.D. Dissertation, Southern Illinois University Carbondale (2011)
[2] Amoroso, E.: Intrusion Detection: An Introduction to Internet Surveillance, Correlation, Trace Back, Traps, and Response. Intrusion.Net Books (1999)
[3] Denning, D.: An Intrusion-Detection Model. IEEE Transactions on Software Engineering 13(2), 118-131 (1986)
[4] Young, C.: Taxonomy of Computer Virus Defense Mechanisms. In : The 10th National Computer Security Conference Proceedings (1987)
[5] Lunt, T.: Automated Audit Trail Analysis and Intrusion Detection: A Survey. In : Proceedings of the 11th National Computer Security Conference, Baltimore, pp.65-73 (1988)
[6] Lunt, T.: A Survey of Intrusion Detection Techniques. Computers and Security 12, 405-418 (1993)
[7] Vaccaro, H., Liepins, G.: Detection of Anomalous Computer Session Activity. In : Proceedings of the 1989 IEEE Symposium on Security and Privacy (1989)
[8] Helman, P., Liepins, G., Richards, W.: Foundations of Intrusion Detection. In : Proceedings of the IEEE Computer Security Foundations Workshop V (1992)
[9] Denault, M., Gritzalis, D., Karagiannis, D., Spirakis, P.: Intrusion Detection: Approach and Performance Issues of the SECURENET System. Computers and Security 13(6), 495-507 (1994)

[10] Forrest, S., Allen, L., Perelson, A., Cherukuri, R.: Self-Nonself Discrimination in a Computer. In : Proceedings of the 1994 IEEE Symposium on Research in Security and Privacy, Los Alamos, CA (1994)
[11] Crosbie, M., Spafford, G.: Defending A Computer System Using Autonomous Agents., COAST Laboratory, Department of Computer Science, Purdue University, West Lafayette, Indiana, USA (1994)
[12] Kumar, S., Spafford, E.: An Application of Pattern Matching in Intrusion Detection., Purdue University (1994)
[13] Ilgun, K., Kemmerer, R., Porras, P.: State Transition Analysis: A Rule-Based Intrusion Detection Approach. IEEE Transactions on Software Engineering 21(3), 181-199 (March 1995)
[14] Esmaili, M., Safavi-Naini, R., Pieprzyk, J.: Evidential Reasoning in Network Intrusion Detection Systems. In : Proceedings of the First Australasian Conference on Information Security and Privacy, pp.253-265 (1996)
[15] Debar, H., Dacier, M., Wespi, A.: Towards a Taxonomy of Intrusion-Detection Systems. Computer Networks 31, 805-822 (1999)
[16] Bace, R.: Intrusion Detection. MacMillan Technical Publishing (2000)
[17] Marin-Blazquez, J., Perez, G.: Intrusion Detection Using a Linguistic Hedged Fuzzy-XCS Classifier System. Soft Computing – A Fusion of Foundations, Methodologies, and Applications 13(3), 273-290 (2008)
[18] Wang, L., Noel, S., et al. Minimum-Cost Network Hardening Using Attack Graphs. Computer Communications 29(18), 3812-3824 (2006)
[19] Dewri, R., Poolsappasit, N., et al. Optimal Security Hardening Using Multi-objective Optimization on Attack Tree Models of Networks. 14th ACM Conference on Computer and Communications Security (2007)
[20] Chen, L. and Leneutre,J. A Game Theoretical Framework on Intrusion Detection in Heterogeneous Networks." IEEE Transactions on Information Forensics and Security 4(2), 165-178 (2009)
[21] Bhargavan, K., Chandra, S., McCann, Peter J. and Gunter, C. A.What packets may come: automata for network monitoring. Proceedings of the 28th ACM SIGPLAN-SIGACT symposium on Principles of programming languages (2001)
[22] Tcpdump/Libpcap: Tcpdump/Libpcap Public Repository. In: Tcpdump.org. Available at: http://www.tcpdump.org/
[23] Chebrolu, S., Abraham, A., Thomas, J.: Feature Deduction and Ensemble Design of Intrusion Detection Systems. Computers and Security 24(4), 295-307 (2005)
[24] Mukkamala, S., Sung, A.: Identifying Significant Features for Network Forensics Analysis Using Artificial Intelligent Techniques. International Journal on Digital Evidence (IJDE) 1(4) (2003)

Adaptive Behaviometric for Information Security and Authentication System using Dynamic Keystroke

Dewi Yanti Liliana
Department of Computer Science
University of Brawijaya
Malang, Indonesia
dewi.liliana@ub.ac.id; dewi.liliana@gmail.com

Dwina Satrinia
Department of Computer Science
University of Brawijaya
Malang, Indonesia
dwina.satrinia@gmail.com

***Abstract*—The increasing number of information systems requires a reliable authentication technique for information security. Password only is not enough to protect user account because it is still vulnerable to any intrusion. Therefore an authentication system using dynamic keystrokes can be the simplest and the best choice. Dynamic Keystroke Authentication System (DKAS) becomes an effective solution which can be easily implemented to gain a high security information system with the aid of a computer keyboard. DKAS verify users based on their typing rythm. Two main stages of DKAS is the enrollment stage to register user into the system, and the authentication stage to check the authenticity of user. Moreover, we use a local threshold to make it becomes adaptive behaviometric for each user. From the experiment conducted, the accuracy rate in distinguishing genuine and impostor user is 91.72%. This shows that the adaptive method of DKAS has a promising result.**

Keywords- authentication system, behaviometric, dynamic keystroke, local threshold

I. INTRODUCTION

The increasing use of information systems in any fields causes a high-demand on a reliable authentication system for information security. Authentication based on biometrics is widely used because of its robustness. Biometrics is a method to recognize human based on intrinsic features or characteristics human has [1]. Physiological biometrics uses unique physical characteristics of individual such as fingerprint, face, palm print, iris, or DNA to identify user and has proven to be a powerful method for authentication systems [1, 2, 3]. Nevertheless, these systems need additional devices (e.g. camera, fingerprint reader, microphone, etc.) to capture human features. Meanwhile, behavioral traits of human or so-called behaviometric which is related to human behavior [4, 5], such as typing rhythm or typing pattern can be implemented on authentication systems without any additional devices. This research implemented behaviometric for authentication system using dynamic keystroke which only needs a computer keyboard to capture the distinct features on typing.

In 2005, Hocquet et.al, conducted a research on authentication system using the combination of password and dynamic keystroke which incorporated three methods; statistical measurement, measure of disorder, and direction similarity measure [5]. The combination method was simple, needed only small size training data, and used global threshold for classifying genuine and impostor users. Global threshold is a constant threshold for all users. The problem was to determine this constant value based on prior knowledge of data. In this research we propose a local threshold setting which can be adaptively adjusted for each different user. Local threshold is adopted from the average score of each user which is obtained during the enrollment phase.

II. DYNAMIC KEYSTROKE AUTHENTICATION SYSTEM

Keystroke means key press. While dynamic keystroke is a biometric which concern about how a user interacts with a keyboard, typing rhythm of a person associated with the habit of typing the password, words, or text [6]. It requires only a keyboard as an input device. Dynamic keystroke also can be implemented for remote access. In addition, biometric based on dynamic keystroke can be used with or without user consciousness.

Password is commonly used on an authentication system for its simplicity, but is less secure because vulnerable to some kinds of attack such as key loggers, spyware, and can be hacked using simple brute force techniques. To enhance the system security and cost efficiency, the password-based authentication system can be combined with dynamic keystroke authentication system (DKAS).

There are two stages on DKAS to distinguish between genuine and impostor user namely, the enrollment stage and the authentication stage (see fig. 1).

At the enrollment stage user sign up their login details such as user name and password which is retyped for several times. The system takes the user dynamic keystrokes ten times for each enrollment, extracts the features, and trains the system to create a reference template of user's typing pattern. The reference template is stored in a database. At the authentication stage, the user enters the login details to be matched with user's reference template which is already stored in the database. This phase consists of collecting user dynamic keystrokes, feature extraction, and feature matching with reference template in the database. The verification process yields two kinds of action: accepted or rejected user access. The first action occurs when the user is the genuine one, while the other action occurs for the impostor user.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

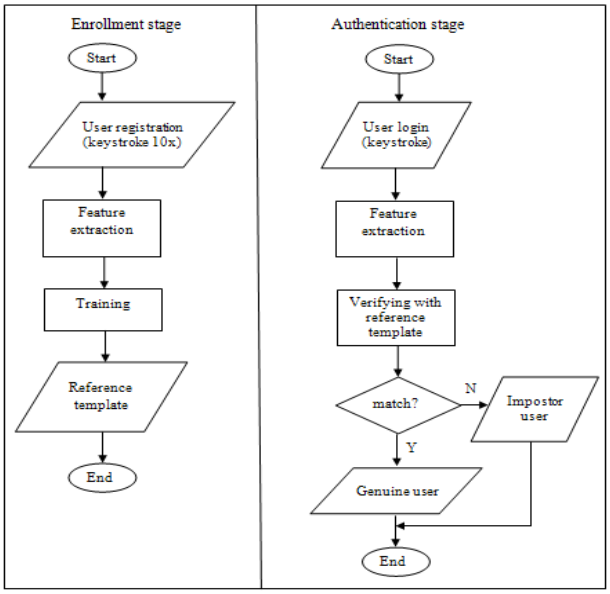


Figure 1. Flowchart of Dynamic Keystroke Authentication System

Four dynamic keystrokes used as features for the authentication system can be seen on illustration of fig. 2.

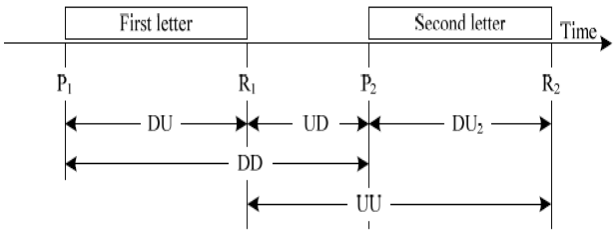


Figure 2. Features of Dynamic Keystroke

Those four features are explained bellow:

1. PP (Press-Press) or DD (down-down) or digraph1: the time between one key press and the next key press (P2-P1).
2. PR (Press-Release) or DU (down-up) or duration: the length of key press (R1-P1).
3. RP (release-press) or UD (Up-down) or latency: the time between key release and the next key press (P2-R1)
4. RR (release-release) or UU (up-up) or digraph2: the time between key release and the next key release (R2-R1).

III. METHODOLOGY

The initial step in this paper is started with the formation of reference templates. Moreover, three methods namely, statistical scoring, measure of disorder, and direction similarity measure will be performed. The last step is the adaptive local threshold setting.

A. The Formation of Reference Templates

In order to verify a user based on dynamic keystrokes, the system needs to create a model or reference template for each user. Reference template is a combination of user keystrokes acquired during the enrollment process which is converted into a more solid form, but still can represent a user keystroke patterns [7]. This research utilized a statistical mean and standard deviation for the reference template formation which can be obtained using equation 1 and 2, respectively.

$$\mu_x = \frac{1}{n}\sum_{i=1}^{n} t_x^i \quad (1)$$

$$\sigma_x = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(t_x^i - \mu_x)^2} \quad (2)$$

where $i=1,2,\ldots,n$ is the number of training samples, $x=1,\ldots,m$ is the number of features used, t_x^i denotes the feature x on the sample i, μ_x and σ_x denote mean and standard deviation of feature x, respectively.

B. Statistical scoring

In the verification process feature matching is performed. It compares the feature of the user test data with the reference template that has been formed on the enrollment stage. Statistical scoring is employed for feature matching. This method will verify the user based on statistical data such as mean and standard deviation. The equation for calculating statistical score is written in Eq.3:

$$Score_{stat} = \frac{1}{n}\sum_{i=1}^{n} e^{-\frac{|t_i-\mu_i|}{\sigma_i}} \quad (3)$$

where $t_i=1,\ldots,n$ is the *i-th* test feature, e is a constant with value of 2.71828, μ_i and σ_i denote mean and standard deviation of reference template vector, respectively.

C. Measure of Disorder

Measure of disorder method is used to compare two ways of typing on the keyboard by studying the similarity between sequences of time features generated as reference templates with sequences of time features which is being tested [8].

To compute the distance between the user keystroke input with the reference template then several steps must be carried out as follows:

1. Rate or rank individual features of each user keystroke input and the comparison reference template. Ordering is done from the smallest to the largest feature value.
2. Calculate the magnitude of differences in rank order or ranking of any existing features on the template with user ratings on keystroke input
3. Calculate the score of disorder using equation 4.

$$score_{disorder} = 1 - \frac{\sum_{i=1}^{n}|R_i^t - R_i^u|}{Max_{disord\ er}} \quad (4)$$

where R_i^t is the *i*-th feature rank obtained from rank vector, R_i^u is the *i*-th feature rank obtained from the user input, and N denotes the number of element or existing features which hold two condition as follows: $Max_{disor\ der} = \frac{N^2}{2}$ if N is even; and

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

$Max_{disorder} = \frac{N^2-1}{2}$ if N is odd.

D. Direction Similarity Measure

Direction similarity measure (DSM) is a simple approach that is discriminatively compares user's typing patterns. The idea of this method is to determine the consistency of the user typing habit. This idea is adopted from the rhythm of the music [8]. In music where the rhythm of a melody is determined by the duration of a tone (the tone is full, half, quarter, etc.), the keystroke is represented by the dynamic rhythm of ups and downs or how quick a keystroke is pressed.

In the calculation of DSM, there is a *ΔD* symbol which is used as a sign of change in the direction of two successive keystrokes. As an example, *ΔD* is positive if there is any time reduction between two keystrokes (faster), and *ΔD* is negative if there is any additional time between two keystrokes (slower). Figure 3 shows the *ΔD* signing.

DU1	DU2	DU3	DU4
245	297	326	268
ΔD :	-1	-1	+1

Figure 3. An example of *ΔD* signing

DSM score can be calculated using the equation 5:

$$Score_{DSM} = \frac{m}{n-1} \quad (5)$$

where *m* is the number of *ΔD* which has the same sign, and *n* is the total features. To compare the user keystroke template with the user keystroke input, what must be considered is the change in sign of *ΔD*. If the sign of *ΔD* from the user reference template equal to the value of *ΔD* of user keystroke input, then the value of *m* increases. The final value of *m* is divided by the number of features minus 1.

E. The incorporation of methods

In this paper the three methods (statistical scoring, measure of disorder, and direction similarity measure) are incorporated by using scoring level which will be done using weighted sum rule operator. The final merged score can be calculated with equation 6:

$$score_{final} = \Sigma(w_i * score^i) \quad (6)$$

where Σw_i=1, score[1] = statistical score; score[2] = measure of disorder score; skor[3] = DSM score.

If the $score_{final}$ of the test user is greater than the user threshold value, then the user will be recognized as a genuine user. Otherwise, it will be recognized as an impostor.

F. Local Threshold

The threshold for the verification system is the similarity value between the test inputs with the model. If the results of feature matching score < threshold, then the user is recognized as an impostor, and if the results of feature matching score ≥ threshold, then the user is recognized as an actual or genuine user.

There are two kinds of threshold, global and local threshold. The global threshold value is set equal to all users, and the local threshold value is set specifically to each user. The problem is to determine the global threshold value required prior knowledge of the data. Therefore, the determination of local threshold value can reduce the problem. Moreover, local threshold can be adaptively adjusted for each different user. There are some ways to estimate local threshold value can be chosen, using the actual user data, impostor data, or a combination of both. The equation used to determine the local threshold value is on Eq. 7:

$$\theta = \mu_{user} - \alpha.\sigma_{user} \quad (7)$$

where θ denotes local threshold, $\mu_{user}, \sigma_{user}$ denotes mean and standard deviation score from user enrollment, respectively, and α denotes a constant factor obtained from the experiment.

The determination of threshold values from user registration data is easy to implement but is less effective because sometimes when the user on registration gets disorders such as drowsiness, talk to or in any uncomfortable situations that are bothering in dynamic keystroke patterns representation. If the threshold was estimated on a situation like this, it will result in decreased accuracy in recognizing user's system. To overcome this problem, we used a method to estimate the weighted scores of local threshold value.

Weighted score is a method to estimate the threshold that gives the weights on the scores based on distance from the user's score to the average score [9]. Scores that were located far from the average are considered as outliers of the user which might be due to a disturbance when users type a password in the registration process. Weighting factor w_i is the parameter of the sigmoid function. w_i values can be calculated by the equation 8:

$$w_i = \frac{1}{1+e^{-C.d_i}} \quad (8)$$

Where *C* is a constant empirically gained from the experiment with the best value = -3. d_i denotes the distance of $score_i$ to the average score ($d_i = |score_i - \mu_{score}|$). Thus, we got the final score S_T by using equation 9:

$$S_T = \frac{\sum_{i=1}^{N} w_i.Score_i}{\sum_{i=1}^{N} w_i} \quad (9)$$

The constant *C* determines the shape of the sigmoid function used to set the weights. $score_i$ and μ_{score} of the training set obtained by a leave-one-out approach. Standard deviation is calculated from $score_i$ against weighted score S_T. The S_T value will replace the μ value of user, and the standard deviation of weighted score will replace the σ user in determining the threshold value. Here are steps on leave-one-out to get $score_i$ value:

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

1. Take a feature vector of n feature vectors used as input during registration for the test.
2. Create a comparison matrix of n-1 remaining feature vectors, then create a reference template of the comparison matrix
3. Compare the test input in step 1 with a reference template that is formed in step 2, using the method used in the verification process to get $score_i$.
4. Repeat steps 1-3 with all possible combinations of the features found on other user registration data so as to produce *n* numbers of $score_i$.
5. Calculate μ_{score} which is an average score of the comparison.

IV. EXPERIMENTS AND RESULTS

Tests carried out using two groups of data that is a typing sample based on user passwords. The first group is users with passwords that usually have been typed by them e.g. their name, etc. The second group is users who use unusual typed words as the password or words chosen at random. Each group consists of the actual and impostor user.

System performance is measured using two error rate: False Rejection Rate (FRR), describes the percentage of a biometric system fails to recognize the actual user and False Acceptance Rate (FAR), describes the percentage of the biometric system identifies incorrect impostor as the actual user. To measure the accuracy of the system, we also measure the Equal Error Rate (EER) obtained when FAR value is equal to FRR (in other words, the intersection of FRR and FAR line). EER is used to compare the performance of different biometric systems [5].

The experiment conducted three kinds of testing: weight value testing that produced the lowest EER value; testing the accuracy of a system that used a local threshold; and testing a system using a global threshold. All tests were using two different groups of data as well as the overall data.

Based on tests done on 826 typed samples, the resulting value of the lowest EER is 8.22%, obtained when the score of statistical weight is 0.7, and the weight score of measure of disorder (MOD) & DSM are 0.15 respectively (see Fig. 4).

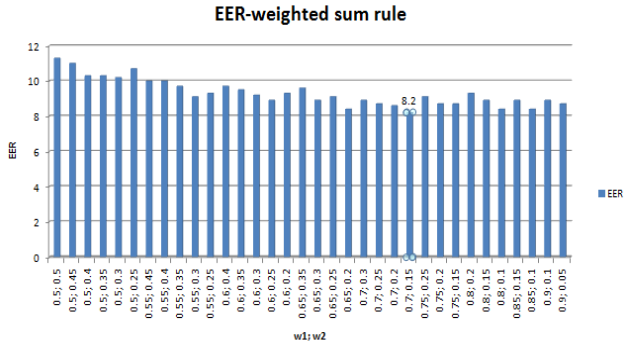


Figure 4. The Equal Error Rate (EER) from the experiment.

The accuracy rate of the authentication system with local and global threshold setting is shown in Table I.

TABLE I. THE ERR COMPARISON OF LOCAL AND GLOBAL THRESHOLD

Data	EER (%)	
	Local	Global
all data	8.22	8
Group 1	4.49	4
Group 2	12	10

From the test result (see table 1), it can be seen that the EER test in group 1 (table 1 row 4) is significantly lower than group 2 (table 1 row 5). This shows that the accuracy rate of dynamic keystroke authentication system depends on the choice of words as passwords. The more accustomed the user with the word, the more the ability of system to recognize users.

From the experiment of comparing global and local thresholds, we got the result which is shown as graphs of error rate in fig. 5. The EER for local threshold is 8.22% with the accuracy rate 91.72%, obtained when the value of α is 1.71. While the EER for global threshold is 8% with the accuracy rate 92%, using the global threshold value = 0.466. When compared with a global threshold, the accuracy rate of a system that uses a local threshold can be said is equally better in verifying the user. The advantages of setting a local threshold is the threshold value for each user can be adaptively estimated using the user data only from the registration process, even without prior knowledge of the data.

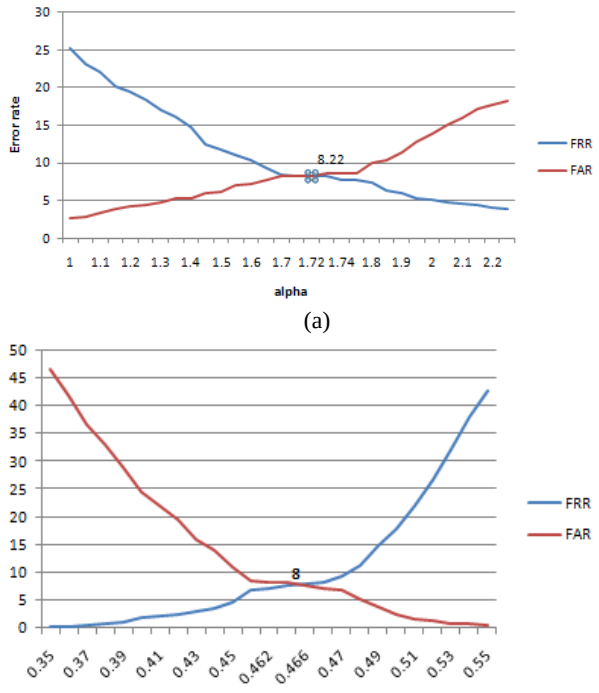


Figure. 5. Graphs of error rate (a) Local Threshold (b) Global Threshold

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

V. CONCLUSION

Dynamic keystroke authentication system is able to verify the user using statistical method, measure of disorder, and direction similarity measure that recognized the user based on the adaptive local threshold. The use of the word or phrase as a password influences the accuracy rate of the system. The accuracy of the system using the local threshold is 91.72%, obtained when the value of α is 1.71.

REFERENCES

[1] N.K. Ratha, J. H. Connell, and R. M. Bolle, "Enhancing security and privacy in biometrics-based authentication systems", IBM systems Journal, vol. 40, pp. 614-634, 2001.

[2] S. Tulyakov, F. Farooq, and V. Govindaraju, "Symmetric Hash Functions for Fingerprint Minutiae", Proc. Int'l Workshop Pattern Recognition for Crime Prevention, Security, and Surveillance, pp. 30-38, 2005.

[3] M.A. Dabbah, W.L. Woo, and S.S. Dlay, "Secure Authentication for Face Recognition", presented at Computational Intelligence in Image and Signal Processing, CIISP 2007, IEEE Symposium, 2007.

[4] http://biosecure.it-sudparis.eu/public_html/biosecure1/public_docs_deli/BioSecure_Deliverable_D10-2-3_b3.pdf

[5] Hocquet, Sylvain, J. Ramel and H. Cardot, "Fusion of Methods for Keystroke Dynamic Authentication", Fourth IEEE workshop on Automatic Identification Advance Technology, 2005.

[6] Hocquet, Sylvain, Jean-Yves Ramel & Hubert Cardot, "User Classification for Keystroke Dynamics Authentication", International Conference on Biometric, Springer-Verlag Berlin Heidelberg. Page 531-539, 2007.

[7] P.S. Teh, B.J.T. Andrew, T. Connie, and S.O. Thian, "Keystroke dynamics in password authentication enhancement", Expert Systems with Application,Vol. 37, Page 8618-8627, 2010.

[8] F. Bergadano, D. Gunetti, and C. Picardi, "User Authentication through Keystroke Dynamics", ACM Transactions on Information and System Security (TISSEC), Page 367-397, New York: ACM New York, 2002.

AUTHORS PROFILE

Dewi Yanti Liliana obtained Bachelor of Informatics from Sepuluh Nopember Institute of Technology Surabaya, Indonesia, in 2004, and Master of Computer Science from University of Indonesia, Depok, Indonesia, in 2009. She is currently working as a Lecturer for the Department of Computer Science, Faculty of Mathematics and Natural Sciences, University of Brawijaya Malang, East java, Indonesia. Her research interests include pattern recognition, biometrics system, computational algorithm, computer vision and image processing.

Dwina Satrinia is a graduate student at the Department of Computer Science, Faculty of Mathematics and Natural Sciences, University of Brawijaya Malang, East java, Indonesia. Her research interests include pattern recognition and biometrics system.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

Denoising Cloud Interference on Landsat Satellite Image Using Discrete Haar Wavelet Transformation

Candra Dewi
Department of Mathematic
University of Brawijaya
Malang, Indonesia
d3w1_c4ndr4@yahoo.com

Mega Satya Ciptaningrum
Department of Mathematic
University of Brawijaya
Malang, Indonesia
megasatya@yahoo.com

Muh Arif Rahman
Department of Mathematic
University of Brawijaya
Malang, Indonesia
arifrahman@ub.ac.id

***Abstract*—Satellite imagery is very useful in information acquisition of the earth's surface image, especially the earth's resources. However, in the process of retrieval information from satellite imagery is often found barriers that can obscure or even cover the imaging of an area. One of these barriers is a cloud, which result the image that covered with lots of noise. Wavelet transformation was usually used to enhance the image or to eliminate striping noise on satellite image. In this paper is used Discrete Haar Wavelet transformation to reduce cloud noise on Landsat TM image. The process includes the Haar Wavelet decomposition of image rows and columns. After that, thresholding process is also applied for de-noising. Thresholding results are then reconstructed using the Inverse Discrete Haar Wavelet. The method is applied to the variation of the band image, the type of thresholding (hard and soft), as well as the size of the image convolution. The testing results on the band 1 to band 6 of Landsat TM imagery showed that the lowest error values are calculated by RMSE (Root Mean Square Error) present in band 1. Image signal to noise ratio in band 1 has the highest value, which means most high-power image signal to noise. This mean that band 1 has the highest pixel value similarity between whole testing data.**

Keywords; Discrete Haar Wavelet, thresholding, image convolution, Landsat TM

I. INTRODUCTION

Image of the earth surface recording can be interpreted by the user for the benefit of various fields. In the process of image acquisition by the satellite, sometimes is found noise that can reduce the image quality. This disorder is caused by the presence of such clouds or fog that can obscure or even covered the satellite during the imaging process [1]. This noise can interfere the interpretation process therefore the results obtained will not be maximal.

Each pixel in the satellite image has some digital value (numeric) in accordance with the band of satellite imagery. For example is Landsat TM image that has 7 bands. Therefore, each pixel has 7 digital values that are suited to 7 band digital value that is owned. The different characteristic each bands causes the difference in the ability to detect clouds. In the study that is performed by Choi and Bindschadler (2004), clouds is very high reflected in the band 2 (0.52 - 0.60 μm).

The elimination process of noise in the spatial domain can be applied directly on image pixels. One of the transformation methods that can be done on the spatial domain is a power-law transformation. While in the frequency domain, the image is broken into multiple kernels to be processed by the analysis of transformation. Transformations that can be done in this domain include Wavelet transformation [3] [4] [5]. Transformation performed to obtain information and identify the original image, by getting its spectrum. Spectrum can be obtained from the image frequency, time, or time-frequency depend on the type of transformation used [6].

It is well known that wavelet transform is a signal processing technique which can display the signals on in both time and frequency domain. Wavelet transform is superior approach to other time-frequency analysis tools because its time scale width of the window can be stretched to match the original signal, especially in image processing studies.

Wavelet transformation can be used to obtain signal both in the frequency domain and time domain. Wavelet time scale width of the window can be stretched to match the original signal. Wavelet is a conversion function that can be used to break up a function or a signal into different frequency components. These components then can be processed in accordance with the scale. While the wave is a function of moving up and down in space and time periodically (sinusoidal), wavelet is a limited wave or sometimes is called as short wave [7].

Haar transform uses the Haar scaling function and the Haar wavelet function. Haar wavelet transformation use the Haar basis functions that is called a wavelet orthonormal [8]. Haar Wavelet functions can be expressed in matrix form.

In the previous study, wavelet transform is used to sharpen the cloud-related shadow areas [1]. Beaulieu et al (2003) refine the resolution of a multi-spectral (MS) image using fusion method and the Stationary Wavelet Transform. In the study performed by Torres and Infante (2001), wavelet transform is used for denoising stripping noise on satellite imagery. This

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

paper applies the Haar wavelet transformation to reduce the noise cloud on Landsat TM imagery.

II. PREVIOUS RESEARCH

The research about the using of wavelet transformation has been done by some researcher. Torres and Infante (2001) present new destriping technique on satellite imagery using Daubechies wavelets of different orders and was tested on a heavily striped Landsat MSS image. Visual inspection and measurement the signal-to-noise ratio showed that the method proved produce encouraging results in image quality and performance, overcome some problems commonly found on traditional destriping techniques and reduce computer time process and storage space.

Beaulieu et al, (2003) refine the resolution of a multi-spectral (MS) image by fusion method using a high-resolution panchromatic (PAN) image and the Stationary Wavelet Transform (SWT). They propose to produce high-resolution MS image that has nearly the same statistical properties than the original multi-spectral image with no blocking image artifacts. These algorithms are based on the injection of high-frequency components from the PAN image into the MS image. They prove that pixel-level fusion was a powerful method to refine the spatial resolution of PAN images.

Wang et al (2003) present a new approach to eliminate the random image noises inherent in the microarray image processing procedure using stationary wavelet transform (SWT) and applied on analysis of gene expression. The testing result on sample microarray images has shown an enhanced image quality. The results also show that it has a superior performance than conventional discrete wavelet transform and widely used adaptive Wiener filter in this procedure.

Elrahman and Elhabiby (2008) developed image sharpening algorithm using wavelet to enhance shadow areas of cloud and tested this algorithm on the panchromatic band of Landsat 7 ETM satellite sub-scenes. The algorithm is applied locally by boosting the image high frequency content in the shadow areas using the defected image de-noised wavelet coefficients. By using visual and quantitative analysis was found that the ability to enhance details under shadow areas increased with the increase in the number of wavelet decomposition levels. Beside, were found that enhancing image quality in the shadow areas could be done using only two or three wavelet decomposition levels.

In these previous studies, the using of wavelets on the satellite image is to sharpen the image and to improve image resolution. Wang et al (2003) already used wavelet to eliminate the noise, but is applied to the gene sequence image. In this paper will be applied discrete Haar wavelet to reduce noise in the form of clouds on satellite images. Although the discrete wavelet transform has a lower performance of the stationary wavelet transform, but its ability to reduce the noise is quite high and does not vary with stationary wavelet transform [5].

III. RESEARCH METHOD

This application was built to reduce noise on Landsat TM satellite image using Haar wavelet transformation method. The limitation of this system includes:

1) The image used is a grayscale image of type TIFF
2) The size of the image used is 256x256 orthonormal

The flowchart of noise reduction process is shown Figure 1. The inputs of this application consist of satellite imagery with clouds noise and image without noise. This input image is presented in grayscale values. Some preprocessing was done to the noise image to reduce the noise. The image without noise is used as a comparison in the testing process.

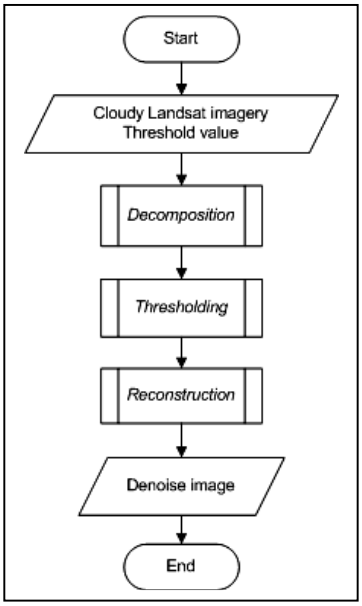


Figure 1 Flowchart of noise reduction process

Firstly, noise image is transformed into the frequency domain using the Haar wavelet transform. The quantization process is then performed using a specific threshold value. The transformation process is performed to the n level, where $N = 2^n$ and N is the size of the image. At each level, the row transformation is done in advance through highpass and lowpass filters. After that, is done transformation of the column.

The next process is tresholding. This process separates pixels based on the degree of grey level values. The wavelet coefficients which are below the threshold are set to zero and than take the other values for purposes of reconstruction of the signal. Threshold used is Hard and Soft Threshold. With ε is the threshold value, hard thresh equation is shown in (1).

$$T_{hard}(X) = \begin{cases} x, & |x| > \varepsilon \\ 0 & |x| \leq \varepsilon \end{cases} \quad (1)$$

On the hard threshold, all wavelet coefficients with a value below a specified threshold are classified as noise and removed (are set to zero). While the coefficients above the

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

threshold is classified as signal. In soft thresholding, the wavelet coefficients with a value below the specified threshold are removed and the wavelet coefficients above the specified threshold are reduced by the threshold value. Thus, this method reduces the range of wavelet coefficients and signal leveling. Soft threshold was chosen because this procedure does not cause non-continuants at x = ± ε. The equation for the soft threshold is shown in (2).

$$T_{soft}(X) = \begin{cases} sign(x)(|x| - thresh), & |x| \geq \varepsilon \\ 0, & |x| < \varepsilon \end{cases} \quad (2)$$

For the determination of threshold values is used equation as in (3).

$$t = \sqrt{\frac{2\sigma^2 \log(n)}{n}} \quad (3)$$

Where:

t = threshold value that is calculated
σ^2 = the variance of data
n = number of data

The equation of variance is shown in (4).

$$\sigma^2 = \frac{\sum (x_i - \bar{x})^2}{(n-1)} \quad (4)$$

The last process is the Inverse Haar Wavelet Transform (IHWT) which is the process of passing the image through the inverse filter matrix transformations. This process is contrary to the decomposition process.

IV. TESTING METHOD

For testing the result is used Root Mean Square Error (RMSE) and Peak-to-Signal Noise Ratio (PNSR).

A. Root Mean Square Error (RMSE)

RMSE is one of the ways to measure the amount of the difference between the estimated values with actual values by measuring the average of error. RMSE is calculated by comparing the number of errors between the denoising image and the original image. The lower the RMSE value the smaller the error calculation has been done. RMSE of digital image with size NxM could be calculated using equation as shown in (5).

$$RMSE = \sqrt{\frac{\sum [f(i,j) - F(i,j)]^2}{N^2}} \quad (5)$$

Where:

f(i,j) is pixel value in original image
F(i,j) is the pixel value on reconstruction image
N^2 is an image size (in pixels)

B. Peak-to-Signal Noise Ratio (PNSR)

PSNR is the comparison between the maximum possible signal strength of a digital signal with the power of noise that affects on the signal (Alfatwa, 2005). PSNR is defined through the signal-to-noise ratio (SNR) to measure the level of signal quality. Signal quality is directly proportional to the value of SNR. The larger of the SNR value, the better the quality of the generated signal. PSNR values usually range between 20 and 40 dB (Alfatwa, 2005). PSNR values can be calculated using equation as shown in (6). Value of 255 represents the upper limit value of image pixels.

$$PSNR = 20 \log_{10}\left(\frac{255}{RMSE}\right) \quad (6)$$

V. SOURCE OF DATA

Image that is used in the testing process is Landsat TM satellite image with each channel has a different sensitivity to the wavelength. Landsat TM satellite orbital period for taking pictures of the earth's surface is generally performed at least 6 months. Satellite imagery from two period of taken picture can be used as a reference on the interpretation process. For example, this study used two images with the same object (the island of Madura) taken in June 2004 and February 2005. In the image taken on 2005 exists cloud covering the particular object and the image taken on 2004 (with the same object) is used as reference.

Preparation of satellite imagery should be done to obtain the image that is suited to analysis. The original image is cropped to the size of 256 x 256 pixels and converted into *Tif* extention format. In addition, the original image with 7 bands is separated per-band for used in applications. Details of the data used are as follows:

1) Landsat image of Madura island, dated June 25, 2004 and dated February 4, 2005
2) Landsat image of Java island, dated June 25, 2004 and dated February 4, 2005

Of the two sources of image data was made 2 pieces of testing data with each of the data contained six band image data (bands 1 to 6 / 7) with each size is 256 x 256 pixels.

Data I: latitude 7:7:45.66 S and longitude 113:3:12. 43 E

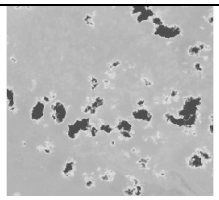
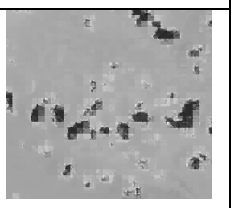
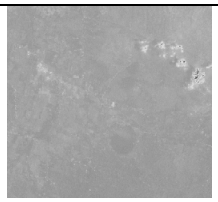
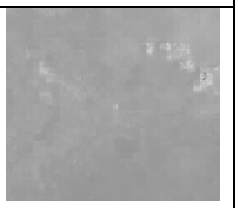
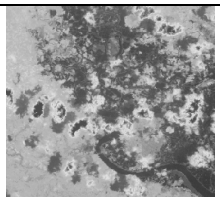
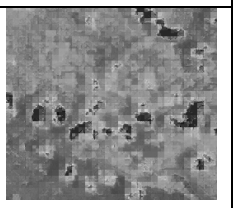
Data II: latitude 7:39:41.99 S and longitude 112:56:41.87 E

VI. RESULT AND DISCUSSION

Some examples of images resulted from denoising process are visually displayed in Table I. The first image shows the result of denoising on data I (band 1) with convolution 2 (hard thresholding), the second on data II (band 1) with convolution 8 (soft thresholding), and the third on data II (band 3) with convolution 8 (hard thresholding).

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

TABLE I. SAMPLE IMAGE OF TESTING RESULT

No	Band	Input Images	Output Images
1	1 (data I)		
2	1 (data II)		
3	3 (data II)		

The RMSE was calculated in the image (the data I and II) which has been transformed with Haar Wavelet. This RMSE values are calculated against several variations of testing which includes testing of inter-thresholding methods, inter-level convolution, and inter-band image. Furthermore, the RMSE is used as input to the calculation of PNSR to observe the ratio of signal strength to noise.

Based on the results in Table 1 could be known that visually processes of Haar wavelet denoising did not show the significant results, because the cloud noise in each band is represented differently. Therefore, an analysis on the basis of testing results on PNSR and RMSE are performed.

The comparison results of RMSE and PNSR on bands 1 to 7 with a convolution of 8, 4, and 2 are shown in Table 2 to Table 7. The RMSE and PNSR are obtained can be used to find out the best *band* on Landsat satellite imagery for the cloud denoising process.

TABLE II. STHE CALCULATION RESULT OF THE RMSE AND PNSR AT CONVOLUTION 8 (DATA I)

Band	Thresh Value	RMSE		PNSR	
		Hard Threshold	Soft Threshold	Hard Threshold	Soft Threshold
1	3,34	30,128	29,633	18,551	18,695
2	2,93	30,98	30,356	18,309	18,486
3	2,93	38,246	37,681	16,479	16,608
4	2,78	42,456	41,998	15,572	15,666
5	2,77	46,93	46,548	14,702	14,773
6/7	2,12	44,859	44,569	15,094	15,15

TABLE III. THE CALCULATION RESULT OF THE RMSE AND PNSR AT CONVOLUTION 8 (DATA II)

Band	Thresh Value	RMSE		PNSR	
		Hard Threshold	Soft Threshold	Hard Threshold	Soft Threshold
1	2,96	7,422	7,041	30,72	31,178
2	2,46	10,339	9,863	27,841	28,251
3	2,06	19,748	19,069	22,22	22,282
4	2,92	43,214	42,623	15,418	15,538
5	2,69	25,312	24,773	20,064	20,251
6/7	1,61	22,551	22,421	21,057	21,118

TABLE IV. THE CALCULATION RESULT OF THE RMSE AND PNSR AT CONVOLUTION 4 (DATA I)

Band	Thresh Value	RMSE		PNSR	
		Hard Threshold	Soft Threshold	Hard Threshold	Soft Threshold
1	3,27	31,306	30,463	18,274	18,455
2	2,86	32,353	31,691	17,932	18,112
3	2,86	39,3	38,646	15,243	16,389
4	2,72	43,3	42,702	15,401	15,522
5	2,71	48,174	47,818	14,475	14,539
6/7	2,08	46,025	45,744	14,871	14,924

TABLE V. THE CALCULATION RESULT OF THE RMSE AND PNSR AT CONVOLUTION 4 (DATA II)

Band	Thresh Value	RMSE		PNSR	
		Hard Threshold	Soft Threshold	Hard Threshold	Soft Threshold
1	2,93	31,988	31,1	18,031	18,276
2	2,57	33,509	32,773	17,628	17,82
3	2,57	40,198	39,393	16,047	16,222
4	2,45	43,854	43,025	15,291	15,456
5	2,45	49,071	48,754	14,314	14,371
6/7	1,89	46,746	46,469	14,736	14,788

TABLE VI. THE CALCULATION RESULT OF THE RMSE AND PNSR AT CONVOLUTION 2 (DATA I)

Band	Thresh Value	RMSE		PNSR	
		Hard Threshold	Soft Threshold	Hard Threshold	Soft Threshold
1	2,89	7,647	7,362	30,461	30,791
2	2,4	10,745	10,262	27,507	27,906
3	2,01	20,123	20,141	22,057	22,049
4	2,86	43,812	43,018	15,299	15,458
5	2,62	26,025	25,486	19,823	20,005
6/7	1,57	22,894	22,873	20,936	20,944

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

TABLE VII. THE CALCULATION RESULT OF THE RMSE AND PNSR AT CONVOLUTION 2 (DATA II)

Band	Thresh Value	RMSE		PNSR	
		Hard Threshold	Soft Threshold	Hard Threshold	Soft Threshold
1	2,59	7,927	7,835	30,148	30,25
2	2,15	11,139	10,667	27,194	27,57
3	1,8	30,564	20,864	21,869	21,75
4	2,56	44,428	43,333	15,178	15,394
5	2,35	26,787	26,279	19,572	19,739
6/7	1,41	23,313	23,478	20,779	20,178

Limitations of different threshold values applied to each image because the distribution of each image pixels values is different. This research calculates the threshold based on the characteristics of image to obtain the best threshold value.

From Table 2, Table 4 and Table 6 can be seen that the band 1 has the smallest RMSE values both in hard thresholding and soft thresholding method (both in the convolution 8, 4, and 2). The lowest RMSE values observed in the convolution 8 with soft thresholding, which is about 29.633 (Data I) and 7.041 (data II). Since the highest RMSE value is detected on band 5 (using hard thresholding with convolution 2) that is about 49.071 (Data I) and 26.787 (Data II). The quite far differences of RMSE value is caused by variations in image value. Data I is an image with a lot of noise distribution, while the data II has less noise in the form of clouds. Base on RMSE value can be seen that band 2 has the lowest error values and band 5 has the highest error value.

The highest value of PNSR is observed on the band I that is around 18.695 dB (data I) and 31.178 dB (data II), while the lowest value is observed in the band 5 with the value is 14.314 dB (data I) and 19.572 dB (data II). It means that the ratio of the image signal to noise at a band I higher than the band 5. The Signal strength value at Data II tends to be higher than the data I, because the noise in the form of clouds fewer than on the Data I. It denoted that the highest probability to perform denoising of cloud can be done on the band 1. On the contrary the lowest probability is on band 5. These results are quite relevant to the characteristic of a band I with a wavelength of 0.45 to 0.52 μm which serves to increase penetration on water body and humidity.

VII. SUMMARY AND CONCLUDING REMARKS

In this paper, Discrete Haar Wavelet methods is applied by utilize thresholding method to the data testing (Landsat satellite image with size 256x256 pixels) to reduce the noise contained in image. The testing result shows that the lowest RMSE value is detected on band 1 (29.633) and highest value is on the band 5 (49.071). As well as the highest PNSR value observe on the band I (18.695 dB) and lowest value is on band 5 (14.314 dB). It can be concluded that the best band to perform denoising clouds with Haar Discrete Wavelet found on the band I, and worst band found on the band 5.

For further study, is proposed to test the result referable reinforced with a system of classification on Landsat satellite imagery.

REFERENCES

[1] A. Abd-Elrahman and M. Elhabiby, “Wavelet Enhancement of Cloud-Related Shadow Areas in Single Landsat Satellite Imagery”, Beijing: The International Archives of the Photogrammetry, Remote Sensing, and Spatial Information Science, Vol. XXXVII part B7, p.1247-1252, 2008.

[2] H. Choi dan R. Bindschadler, “Cloud Detection in Landsat Imagery of Ice Sheets Using Shadow Matching Technique and Automatic Normalized difference Snow Index Threshold Value Decision”, Remote Sensing of Environment, Vol. 91. p.237-242, 2004.

[3] J. Torres and S.O. Infante, “Wavelet Analysis for The Elimination of Striping Noise In Satellite Images”, Society of Photo-Optical Instrumentation Engineers, DOI: 10.1117/1.1383996, 2001.

[4] M. Beaulieu, M., S. Faucher, and L. Gagnon, « Multi-Spectral Image Resolution Refinement Using Stationary Wavelet Transform”, International Geoscience Remote Sensing Symposium, Vol. 6, pp. 4032–4034, 2003.

[5] X.H. Wang, Robert S.H. Istepanian, and Y.H. Song, “Microarray Image Enhancement By Denoising Using Stationary Wavelet Transform”, IEEE Transactions on Nanobioscience, Vol. 2, No.4, 2003.

[6] D. F. Alfatwa, “Watermarking pada Citra Digital Menggunakan Discrete Wavelet Transform”, Informatic Study Program, Technoly Institute of Bandung, 2005.

[7] R. B. Edy Wibowo, “Scattering Problem for A System of Non Linear Klein-Gordon Equations Related to Dirac-Klein-Gordon Equations”, An International Multidisciplinary Journal, Vol. 71, No. 3-4, 2009.

[8] Gonzales, Rafael C dan Woods, Richard E. 2005. Digital Image Processing, 2nd Edition. New Jersey: Prentice Hall

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

Calculating Rank of Nodes in Decentralised Systems from Random Walks and Network Parameters

Sunantha Sodsee* † ‡, Phayung Meesad*, Mario Kubek†, Herwig Unger†
*King Mongkut's University of Technology North Bangkok, Thailand
†Fernuniversität in Hagen, Germany
‡ Email: sunantha.sodsee@fernuni-hagen.de

***Abstract*—To use the structure of networks for identifying the importance of nodes in peer-to-peer networks, a distributed link-based ranking of nodes is presented. Its aim is to calculate the nodes' PageRank by utilising the local-link knowledge of neighborhood nodes rather than the entire network structure. Thereby, an algorithm to determine the extended PageRank, which is called NodeRank of nodes by distributed random walks that supports dynamic P2P networks is presented here. It takes into account not only the probabilities of nodes to be visited by a set of random walkers but also network parameters as the available bandwidth. NodeRanks calculated this way are then applied for content distribution purposes. The algorithm is validated by numerical simulations. The results show that the nodes suited best to place sharable contents in the community on are the ones with high NodeRanks, which also offer high-bandwidth connectivity.**

***Index Terms*—Peer-to-peer systems, PageRank, NodeRank, random walks, network parameters, content distribution.**

I. Introduction

At present, the amount of data available in the *World Wide Web* (WWW) is growing rapidly. To ease searching for information, several web search engines were designed, which determine the relevance of keywords characterising the content of web pages and return all search results to querying users (or nodes) such as an ordinary index-based keyword search method. Usually, there are more results than users are expecting and able to handle. As a consequence of this, a ranking of query results is needed to facilitate searchers to access lists of search results ranked according to keyword relevance.

In particular, the search engine *Google* is based on keywords. To improve its search quality, a link analysis algorithm called *PageRank* [1] is used to define a rank of any page by considering the page's linkage. The importance of a web page is assumed to correlate to the importance of the pages pointing to it. Another link-based algorithm is the *Hyperlink-Induced Topic Search (HITS)* [2]. It maintains a hub and authority score for each page, in which the authority and hub scores are computed by the linkage relationship of pages. Both *PageRank* and *HITS* have an ability to determine the rank of keyword relevance but they are iterative algorithms. These algorithms require centralised servers, since they process knowledge on the entire Internet. Consequently, they cannot be applied in decentralised systems like peer-to-peer (P2P) networks.

Because of its higher fault tolerance, autonomy, resource aggregation and dynamism, the content-based presentation of information in P2P networks has more benefits than the traditional client-server model. One of the crucial criteria for the use of the P2P paradigm is the search effectiveness made possible. The usually employed search method based on flooding[4] works by broadcasting query messages hop-by-hop across networks. This approach is simple, but not efficient in terms of network bandwidth utilisation. Another method, distributed hash tables based search (DHT) [3] is efficient in terms of network bandwidth, but causes considerable overhead with respect to index files. DHT does not adapt to dynamic networks and dynamic content stored in nodes. Exhibiting fault tolerance, self-organisation and low overhead associated with node creation and removal, conducting *random walks* is a popular alternative to flooding [5]. Many search approaches in distributed search systems seek to optimise search performance. The objective of a search mechanism is to successfully return desired information to a querying user. In order to meet this goal, several approaches, e.g. [5], [6], were proposed. Most of them, however, base search on content, only.

Due to the efficiency of [1] in the most-used search engine, the link analysis algorithm PageRank for determining the importance of nodes has become a significance technique integrated in distributed search systems as it is not only sensible to apply it in centralized system for improving query results, but can also be of use in distributed systems. [7], [8] and [9] proposed distributed PageRank computations. The work in [7] is based on iterative aggregation-disaggregation methods. Each node calculates a PageRank vector for its local nodes by using links within sites. The local PageRank will be updated by communicating with a given coordinator. For [8] and [9], nodes compute their PageRank locally by communicating with linked nodes. Moreover, [9] presented that each node exchanges its PageRank with nodes to which it links to and those linking to it and paid attention to only parts of the linked nodes required to be contacted. Nevertheless, the mentioned works do not employ any network parameters in defining PageRank, which could be of advantage to reduce user access times.

Herein, the first contribution of this paper is to introduce an improved notion of PageRank applied in P2P networks which works in a distributed manner. When conducting searches, not

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

only matching content but also content accessibility is considered which will influence the rank calculations presented. Therefore, a distributed algorithm based on random walks is proposed which takes network parameters, of which bandwidth is the most important one, into consideration when calculating ranks, which is called NodeRank. This novel NodeRank determination will be described in Sec. III, after the state of the art has been outlined in Sec. II. The second contribution is to enhance the search performance in hybrid P2P systems. The presented NodeRank formula can be applied not only to support information retrieval but also content distribution in order to find the most suitable location for contents to be distributed. Contents will be distributed by artificial behavior of random walkers, which is based on a modified ant-based clustering algorithms, to pick from specific nodes and place contents on the most suitable location based on the presented NodeRank definition. Its details will be presented in Sec. IV.

II. State of the Art

In this section, the background of P2P systems is presented first. Then, ant-based clustering algorithms are introduced. Later, the PageRank formula according to [1] is described. Finally, the simulation tool *P2PNetSim* used in this work is presented.

A. P2P Systems

Currently, most of the traffic growth in the Internet is caused by P2P applications. The P2P paradigm allows a group of computer users (employing the same networking software) to connect with each other to share resources. Peers provide their resources such as processing power, disk storage, network bandwidth and files to be directly available to other peers. They behave in a distributed manner without a central server. As peers can act as both server and client then they are also called *servent*, which is different from the traditional client-server model. In addition, P2P systems are adaptive network structures whose nodes can join and leave them autonomously. Self-organisation, fault-tolerance, load balancing mechanisms and the ability to use large amounts of resources constitute further advantages of P2P systems.

1) System Architectures: At present, there are three-major architectures for P2P systems, viz. unstructured, hybrid and structured ones.

In unstructured P2P systems, however, such as Gnutella [4], a node queries its neighbours (and the network) by flooding with broadcasts. Unstructuredness supports dynamicity of networks, and allows nodes to be added or removed at any time. These systems have no central index, but they are scalable, because flooding is limited by the messages' time-to-live (TTL). Moreover, they allow for keyword search, but cannot guarantee a certain search performance.

Cluster-based hybrid P2P systems or hybrid P2P systems are a combination of fully centralised and pure P2P systems. Clustering represents the small-world concept [15], because similar things are kept close together, and long distance links are added. The concept allows fast access to locations in searching. The most popular example for them is KaZaA [13]. It includes features both from the centralized sever model and the P2P model. To cluster nodes certain criteria are used. Nodes with high storage and computing capacities are selected as *super nodes*. The normal nodes (*clients*) are connected to the super nodes. The super nodes communicate with each other via inter-cluster networks. In contrast, clients within the same cluster are connected to a central node. The super nodes carry out query routing, indexing and data search on behalf of the less powerful nodes. Hybrid P2P systems provide better scalability than centralised systems, and show lower transmission latency (i.e. shorter network paths) than unstructured P2P systems.

In structured P2P systems, peers or resources are placed at specified locations based on specific topological criteria and algorithmic aspects facilitating search. They typically use distributed hash table-based indexing [3]. Structured P2P systems have the form of self-organising overlay networks, and support node insertion and route look-up in a bounded number of hops. Chord [10], CAN[11] and Pastry [12] are examples of such systems. Their features are load balancing, fault-tolerance, scalability, availability and decentralisation.

2) Search Methods: Generally, in P2P systems, three kinds of content search methods are supported. First, when searching with a specific keyword, the query message from the requesting node is repeatedly routed and forwarded to other nodes in order to look for the desired information. Secondly, for advertisement-based search [14], each node advertises its content by delivering advertisements and selectively storing interesting advertisements received from other nodes. Each node can locate the nodes with certain content by looking up its local advertisement repository. Thus, it can obtain such content by a one-hop search with modest search cost. Finally, for cluster-based search, nodes are grouped according to the similarity of their contents in clusters. When a client submits a query to a server, it is transmitted to all nodes whose addresses are kept by the server, and which may be able to provide resources possibly satisfying the query's search criteria.

In this paper, cluster-based P2P systems are considered in the example application, which combines the advantages of both the centralised server model and distributed systems to enhance search performance.

B. Ant-based Clustering Methods

In distributed search systems, data clustering is an established technique for improving quality not only in information retrieval but also distribution of contents. Clustering algorithms, in particular ant-based ones, are self-organizing methods -there is no central control- and also work efficiently in distributed systems.

Natural ants are social insects. They use a stigmergy [16] as an indirect way of co-ordination between them or their actions. This gave rise to a form of self-organisation, producing intelligence structures without any plans, controls or direct communication between the ants. Imitating the behaviour of ant societies was first proposed to solve optimisation problems by *Dorigo* [17].

In addition, ants can help each other to form piles of items such as corpses, larvae or grains of sand by using the

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

stigmergy. Initially, ants deposit items at random locations. When other ants visit the locations and perceive deposited items, they are stimulated to deposit items next to them. This example corresponds to cluster building in distributed computer networks.

In 1990, *Deneubourg et al.* [18] first proposed a clustering and sorting algorithm mimicking ant behaviour. This algorithm is implemented based on corpse clustering and larval sorting of ants. In this context, clusters are collections of items piled by ants, and sorting is performed by distinguishing items by ants which place them at certain locations according to item attributes. According to [18], isolated items should be placed at locations of similar items of matching type, or taken away otherwise. Thus, ants can pick up, carry and deposit items depending on associated probabilities. Moreover, ants may have the ability to remember the types of items seen within particular durations and moved randomly on spatial grids.

Few years later, *Lumer* and *Faieta* [19] proposed several modifications to the work above for application in data analysis. One of their ideas is a similarity definition. They use a distance such as a Euclidean one to identify similarity or dissimilarity between items. An area of local neighbourhood at which ants are usually centered is defined. Another idea suggested for ant behaviour is to assume short-term memory. An ant can remember the last m items picked up and the locations where they have been placed.

The above mentioned contributions pioneer the area of ant-based clustering. At present, the well-known ant-based clustering algorithms are being generalised, e.g. in *Merelo* [20].

C. The PageRank Algorithm

As in hybrid P2P architectures, good locations of clusters can improve search performance. To find suitable locations, ranking algorithms can be applied.

Herein, the *PageRank* (PR) algorithm, introduced by Brin and Page [1], is presented that is well-known, efficient and supports networks of large sizes. Based on link analysis, it is a method to rank the importance of based on incoming links. The basic idea of *PageRank* is that a page's rank correlates to the number of incoming links from other, more important pages. In addition, a page linked with an important page is also important [7]. Most popular search engines such as *Google* employ the *PageRank* algorithm to rank search results. *PageRank* is further based on user behaviour: a user visits a web page following a hyperlink with a certain probability η, or jumps randomly to a page with probability $1-\eta$. The rank of a page correlates to the number of visiting users.

Classically, for PageRank calculation the whole network graph needs to be considered. Let i represent a web page, and J be the set of pages pointing to page i. Further, let the users follow links with a certain probability η (often called damping factor) and jump to random pages with probability $1-\eta$. Then, with the out-degree $|N_j|$ of page j PageRank PR_i of page i is defined as

$$PR_i = (1-\eta) + \eta \sum_{j \in J} \frac{PR_j}{|N_j|}. \qquad (1)$$

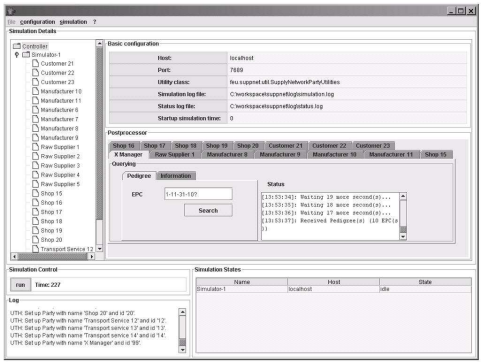


Fig. 1. *P2PNetSim*: simulation tool for large P2P networks

The damping factor η is empirically determined to be $\approx 90\%$.

D. The Simulation Tool P2PNetSim

The modified PageRank calculation presented here will be considered in general setting. In order to carry out experiments, the conditions of real networks are simulated by using the artificial environment of the distributed network simulator *P2PNetSim* [21]. This tool was developed, because neither network simulators nor other existing simulation tools are able to investigate, in decentralised systems, processes programmed on the application level, but executed in real TCP/IP-based network systems. This means, a network simulator was needed that is capable of

- simulating a TCP/IP network with an IP address space, limited bandwidth and latencies giving developers the possibility to structure the nodes into subnets like in existing IPv4 networks,
- building up any underlying hardware structure and establishing variable time-dependent background traffic,
- setting up an initial small-world structure in peer neighbourhood warehouses and
- setting up peer structures allowing the programmer to concentrate on programming P2P functionality and to use libraries of standard P2P functions like broadcasts.

Fig. 1 presents the simulation window of *P2PNetSim*. The simulator allows to simulate large-scale networks and to analyse them on cluster computers, i.e. up to 2 million peers can be simulated on up to 256 computers. The behaviour of all nodes can be implemented in Java and, then, be distributed over the nodes of the network simulated.

At start up, an interconnection of the peers according to the small-world concept is established in order to simulate the typical physical structure of computers connected to the Internet. *P2PNetSim* can be used through its graphical user interface (GUI) allowing to set up, run and control simulations. For this task, one or more simulators can be set up. Each simulator takes care of one class A IP subnet, and all peers within this subnet. Each simulator is bound to a so-called simulation node, which is a simulator's execution engine. Simulation nodes reside on different machines and, therefore, work in parallel. Communication between peers within one subnet is confined to the corresponding simulation node. This

hierarchical structure, which is based on the architecture of real IP networks, provides *P2PNetSim* with high scalability. *P2PNetSim* is based on Java. Users can implement their own peers for simulation just by writing Java programs that inherit from the *P2PNetSim* peer class. These peers provide basic communication and logging facilities as well as an event system which allows tracking of the state of simulation and to perform analysis processes. Due to its applicability for large-scale P2P networks simulations, *P2PNetSim* is utilised to simulate the performance of the presented work.

III. Modified Rank of Nodes Calculation

As the first contribution of this paper, in the present section an algorithm for the calculation of PageRanks in a modified way is presented. PageRanks are calculated in decentralised systems in the course of random walks. A new method to apply the algorithm incorporating network parameters will be introduced later.

A. Basic Ideas

The PageRank of a node in a network can also be represented as the node's probability to be visited in the course of a random walk through the network. If the node is visited many times by random walkers, then the node is assumed to be more important than the less often visited ones. Random walks require no knowledge of network structure, and are attractive to be applied in large-scale dynamic P2P networks, because they use local up-to-date information, only. Moreover, they can easily manage connections and disconnections occurring in networks. Their shortcoming, however, is time consumption, especially in the case of large networks [22]. To address this problem, it is proposed to utilise a set of random walks carried out in parallel. The first objective here is to prove that the performance of determining PageRanks with this approach is equivalent to the one of *PageRank* [1].

In addition to random walks, also network parameters shall be incorporated into PageRank calculations. In this context, the bandwidth of communication links is the most important parameter. Consequently, capacity figures must influence the PageRank formula. The transition probability characteristic for random walks will also be considered. Random walkers move to any of a node's neighbours with non-equal probabilities [23] depending on the network capacities. The second objective here is to show the performance of the modified PageRank calculation under the influence of network parameters.

B. PageRank Definition by Random Walking

Let $G = (V, E)$ be an undirected graph to represent network topologies, where V is the set of nodes $v_i,\ i = \{1, 2, \ldots, n\}$, and $E = V \times V$ is the set of links e_{ij} and n is the number of nodes in the network. In addition, the neighbourhood of node i is defined as $N_i = \{v_j \in V | e_{ij} \in E\}$.

Typically, a random walker on G starts at any node v_i at a time step $t = 0$. At $t = t + 1$, it moves to $v_j \in N_i$ selected randomly with a uniform probability p_{ij}, where $p_{ij} = \frac{1}{|N_i|}$ is the transition probability of the random walker to move from v_i to v_j in one step.

The importance of a node, given by its *PageRank*, at time $t > 0$ is defined as the number of times that random walkers have visited the node so far: $PR_i(t) = \frac{f_i(t)}{step(t)}$. Note that $\sum_i PR_i(t) = 1$ when $t \to \infty$, where $f_i(t)$ is the number of visits to v_i and $step(t)$ its number of steps up to time t, respectively.

If the number of random walkers is increased to $k \in \mathbb{N}$, then the PageRank can be calculated by

$$PR_i(t) = \frac{f_{i_k}(t)}{\sum_k step_k(t)}, \tag{2}$$

where $f_{i_k}(t)$ is the number of all k random walkers' visits taken place so far in the $step_k(t)$ steps until time t.

The PageRank of the whole network can be defined as the average PageRank:

$$\overline{PR} = \frac{\sum_i PR_i}{n} = \frac{1}{n}. \tag{3}$$

In fact, due to dynamicity, the exact network size n cannot be known in distributed systems. Hence, to calculate the average PageRank, n is estimated as

$$n = \frac{\sum_i PR_i}{\overline{PR}} = \frac{1}{\overline{PR}}. \tag{4}$$

In other words, the network size is estimated from a sample of PR values whose mean value will converge to $\frac{1}{n}$.

C. Influence of Network Parameters on Transition Probability

To study the influence network parameters have on the importance of nodes, the bandwidth of communication links shall be applied here to identify -generally non-uniform- transition probabilities of random walkers, i.e. if a node is connected by a low-bandwidth link, then the probability to be reached will be lower than via a high-bandwidth one. Herein, the NodeRank is introduced.

Let $B(e_{ij})$ be the bandwidth of the link connecting nodes v_i and v_j. Then, the transition probability of random walkers to move from v_i to v_j is defined as

$$p_{ij} = \frac{B(e_{ij})}{\sum_{j \in |N_i|} B(e_{ij})}, \tag{5}$$

where $\sum_{j \in N_i} p_{ij} = 1$. The number of times that random walkers have visited the node $f_{i_k}(t)$ influences the visiting probability of the random walkers and the NodeRank (NR) is calculated by

$$NR_i(t) = \frac{f_{i_k}(t)}{\sum_k step_k(t)}. \tag{6}$$

Eq. 5 can also be applied when further network parameters are taken into consideration by replacing $B(e_{ij})$ by other quantities or combining it with other parameters.

D. Convergence Behaviour

In this subsection, the convergence behaviour of PageRank values determined by random walks is studied. Convergence time is defined as the duration until a probability, stable within a certain margin, of being visited is reached by all nodes.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

This usually small margin ϵ [8] is defined as the maximum PageRank values may change between two time subsequent steps. Convergence is reached when $|PR_i(t)-PR_i(t-1)| \leq \epsilon$ is fulfilled for all nodes.

In order to avoid the chaotic vary of PageRank values, a mean value (rather than 0) is identified to be an initial PageRank of nodes. The final PageRank values can be more or less than the initial ones, then they will be changed smoothly. Then, the PageRank is calculated as

$$PR_i(t) = \frac{1}{n}e^{-ct} + \frac{f_{i_k}(t)}{\sum_k step_k(t)}(1 - e^{-ct}), \qquad (7)$$

where n is the estimated number of nodes in the network, c is a damping factor, $f_{i_k}(t)$ is the number of the random walkers' visits to v_i after $step_k(t)$ steps until time t. As a first estimation, the term $\frac{1}{n}e^{-ct}$ represents the initial value assigned to the PageRank. For $t = 0$, this term e^{-ct} assumes the value 1, $1-e^{-ct}$ vanishes and, thus, the initial PageRank of all nodes becomes $PR_i(0) = \frac{1}{n}$. On the other hand, for $t \to \infty$, e^{-ct} vanishes, $1 - e^{-ct}$ approaches 1 and the PageRank assumes the same value as in Eq. 2, viz. $PR_i(t) = \frac{f_{i_k}(t)}{\sum_k step_k(t)}$. In this case, the PageRank calculations of all nodes start with the same initial value, the parameter c may range within $0 < c < 1$ and its value also effects the convergence time.

E. Comparative Evaluation

The objective pursued in this subsection is an empirical proof of concept. The following issues are addressed:

1) Is the PageRank generated by sets of random walks equivalent to the one rendered by the algorithm of Page and Brin?
2) Can the average PageRank of a network be estimated by considering only a part of the network and, if so, which size does this network need to have?
3) How long is the convergence time, and how does it depend on network size, network structures and number of random walks?
4) How do network parameters influence NodeRank?

Due to reliability, toleration of the node's failure and no redundancy of connection, hereby, the proof is simulated on grid-like overlay network structures, which are a grid and a torus. For the grid structure, the maximum degree of a node is four and a minimum one is two. In contrast, a degree of all nodes is four for the toroidal grid structure. The sizes of networks are represented as the multiplication between the number of x-columns and y-rows, and a node is represented by a cross between x-columns and y-rows.

1) Generating PageRank by Sets of Random Walks: To conduct comparative simulations, a rectangular network (or grid) with the size of 20×20 was used and the margin ϵ selected as 8×10^{-7}. First, considering the *PageRank* algorithm, Eq. 1 was applied. At time $t = 0$, the PageRank of all nodes was set to an initial value. Each node calculated its PageRank and, then, distributed its updated PageRank to its set of neighbours N_i. At every time step, the updated PageRank was compared with the previous one. If their difference turned out to be below the margin ϵ, the obtained value was regarded as the node's PageRank. On the other hand, to investigate the calculation of PageRanks based on k random walkers, Eq. 2 was considered and k selected as 50. The random walkers visited nodes until t=120, 946, then convergence of PageRank values was reached. The results obtained for both approaches are shown in Fig. 2. Due to the structure of the grid, the PageRank of a node depended on its number of links. The node that had the lowest number of links had the lowest PageRank too. Consequently, the results revealed that a set of random walks produced the same PageRank as the algorithm *PageRank* of Page and Brin.

2) Approximating Average PageRank: In this subsection it will be shown that by calculating an average PageRank it is possible to estimate the size of P2P networks, which is generally not known.

For this purpose, simulations were conducted on grids with the size of 20×20 and 50×50, respectively, and by using $k = 50$ random walkers, yielding as exact average PageRank $\overline{PR} = 2.5 \times 10^{-3}$ and $\overline{PR} = 4 \times 10^{-4}$, respectively.

For both simulations, only fractions of the networks were queried, with the fraction sizes ranging from just a small number of nodes to around 80% of the overall network size. Calculating mean PageRanks from these data indicated that they were close to the exact average PageRank values, which could be proved for fractions with a tenth of the networks' size or larger.

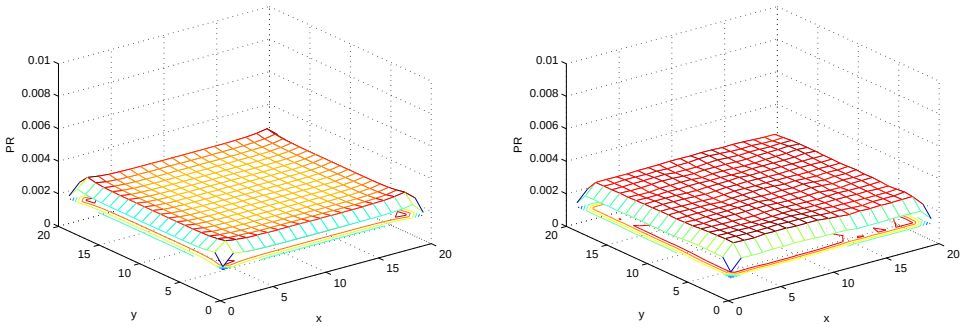
The simulation was started by sampling the PageRank values from 50 nodes (it was 0.2% of network size) and went on until taking 2,000 nodes (it was 80% of network size) into consideration. The approximate average PageRank reached the exact value with a deviation of just 4×10^{-4} already by 250 nodes or more.

To conclude, if the sample size of nodes would be large enough to calculate the approximate $\overline{PR}$, then this value could be used to estimate the network size $n = \frac{1}{\overline{PR}}$.

3) Convergence of PageRank Determination by Random Walking: Convergence behaviour was studied based on three experiments. In the first one, the convergence time for a single walker was compared for different network sizes. Here, simulations in both grid and toroidal grid structures were conducted with the margin $\epsilon = 0.0001$. The number of nodes (n) was increased from small to large network size, and set to 100, 400, 900, 1,600, 2,500 and 10,000, respectively. In the simulations, n represented the network size, while in real networks one has to settle for an estimated value. For $\epsilon = 0.0001$ and the toroidal grid, random walks led to faster convergence than for the grid structure especially when the number of nodes exceeded 1,600. In addition, for both grid and toroidal grid, random walks in small networks led to earlier convergence than the bigger ones.

In the second experiment, the number of random walkers was increased to $k = 50$ in order to save time by parallel processing. Its convergence time was compared to the one obtained for single random walker. Here, both a grid and a toroidal grid with 20×20 nodes and the very small $\epsilon = 8 \times 10^{-7}$ were used. The results show that convergence was $\approx$ 45–50 times slower for single random walker than for the fifty walkers working in parallel, for both network structures

http://sites.google.com/site/ijcsis/
ISSN 1947-5500



(a) Ranking with *PageRank* (b) Ranking with random walks

Fig. 2. Comparison of ranking on a grid with size 20×20 ($\epsilon = 8 \times 10^{-7}$)

TABLE I
CONVERGENCE TIMES FOR DIFFERENT NUMBERS OF RANDOM WALKERS
($n = 400$, $\epsilon = 8 \times 10^{-7}$)

Walkers	**Grid**	**Toroidal Grid**
	$c = 0.401 \times 10^{-3}$	$c = 0.401 \times 10^{-3}$
1	$7,011,870$	$6,164,214$
10	$683,811$	$640,994$
20	$337,850$	$295,375$
50	$123,990$	$115,284$

considered. From this simulation it could be concluded that the number of random walkers effected the convergence time at $\epsilon = 8 \times 10^{-7}$. If ϵ was very small, here it turned out that random walks in the grid reached convergence slower than in the torus.

In the third experiment the influence of the damping factor c was studied. Again, a grid and a toroidal grid with 400 nodes were considered. The margin ϵ was selected as 8×10^{-7}) and the number of random walkers increased to be $k = \{1, 10, 20, 50\}$, respectively. The simulation results for both network structures revealed that a suitable value for c value was important according to Eq. 7. If c was, for instance, too small, i.e. $c \leq 0.4 \times 10^{-3}$, then Eq. 7 would not support PageRanking. The suitability of c values was determined by ϵ value and the number of nodes. For $n = 400$ and $\epsilon = 8 \times 10^{-7}$ suitable values for c were slightly greater than 0.4×10^{-3}. The convergence times for both grid and torus are given in Table I. It showed that c and the number of random walkers effected the convergence time for both structures.

4) Considering Link Bandwidths: In this subsection, the bandwidth of communication links is taken into account. Users of P2P networks may use various link bandwidths available. Consequently, node accessibility is also different. Herein, for a high bandwidth the data transfer rate is assumed to be 100 Mbps, in contrast, 30 Mbps is supposed to be a low rate one, which is around three times slower than the high bandwidth one.

The simulations considering the link bandwidths were carried out in the same settings as above, viz. 20×20 and 50×50 nodes in both a grid and a torus, with 50 random walkers and $\epsilon = 8 \times 10^{-7}$. The effect of varying link bandwidths is shown in Fig. 3. The results showed that NodeRanks were influenced by the bandwidth of communication links in such a way that the probability of a node being visited by random walkers correlated to the bandwidth of the links leading to it. Hence, NodeRanks depended on link bandwidths. In other words, a node connected by high-bandwidth links will be more important than a node with the same topological properties, but connected to lower-bandwidth links.

IV. A REAL-WORLD EXAMPLE: CONTENT DISTRIBUTION

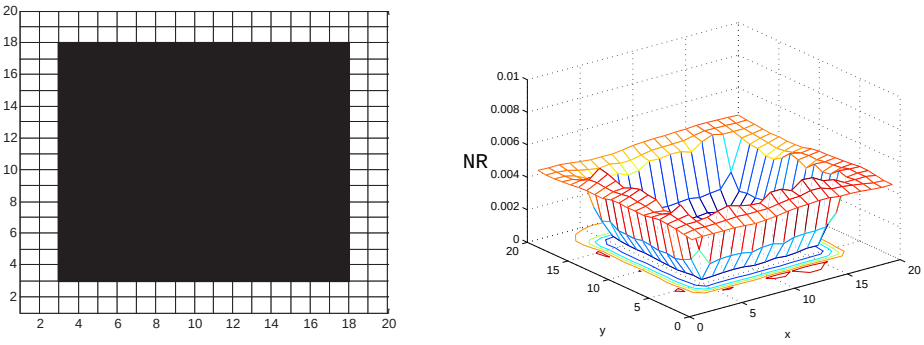
In this section, the second contribution of the paper is presented, showing that the NodeRank as defined here can also be applied to content distribution networks.

A. Introduction

As mentioned in Sec.I, client-server application models are not suitable anymore to serve contents of high demand such as audio and video files and software packages. Typically, a content provider utilises centralised servers, which often suffer from congestion and slow network speed when the demand for the provided content increases. Therefore, content distribution techniques are deployed [24], where content is delivered to a large number of clients through surrogate servers that hold copies from the original server to reduce its load as well as to improve end-user performance, and increase global availability of contents. When a client tries to access contents, the respective query is routed to the surrogate server closest to the client in order to speed up the delivery of contents.

Especially video-on-demand (VoD) services, which play an increasing role in businesses and in education, have to handle a large amount of data and therefore should employ content distribution techniques. This is especially true since VoD services additionally must fulfil low latency constraints [28], allow random frame access and seeking to provide a user experience on the same level of quality as known from local file playback. Due to their inherent scalability, P2P-based approaches can overcome the disadvantages of client-server based architectures, since each peer can act as streaming client and server at the same time. Cluster-based hybrid P2P systems are considered as solutions which combine the advantages of P2P technologies and client-server models [29].

http://sites.google.com/site/ijcsis/
ISSN 1947-5500



(a) Link bandwidths in a torus with a size 20×20 (b) NodeRanks for the torus shown left

Fig. 3. NodeRanks determined by fifty random walks for substructures of different bandwidths ($\epsilon = 8 \times 10^{-7}$)

Remark: To ease visualization, link bandwidths are determined in an area: high-bandwidth links area and low-bandwidth links area. The high-bandwidth links area consists of fade-black lines and dark-black ones denote low-bandwidth links area.

The suitable location for video files in cluster-based hybrid P2P networks should be based on three major factors effecting the performance of P2P systems: contents, network parameters and user behavior. The video files should be placed on nodes as follows:

1) Nodes with a central position.
2) Nodes with high speed and low-latency network connections, which support the above mentioned quality of service requirements.
3) Nodes, which close to those users, who frequently access files (in this paper, the third factor -user behavior- is not considered yet).

Existing solutions for cluster-based hybrid P2P networks can be characterised by robustness and high service availabliliy. Their drawbacks, however, are (a) high network traffic caused by routing and replication, and (b) the necessary consistency management of multiple copies of data. To avoid these issues, a community concept is considered in this article. In the proposed network model, nodes are grouped in a community based on their interest. Contents should be distributed to a known node with high bandwidth in a community, known as a super node in cluster-based hybrid P2P systems, in order to combine the advantages of the client-server model and pure P2P systems (refer to Sec. II-A). The super node is responsible for maintaining the contents stored on it. Content updates can be performed by the respective content's owner and will be propagated to locations of replicated copies. When searching for contents, user nodes or clients in the community will send queries directly to the super node and therefore reduce the network traffic because a specific content does not need to be placed on many locations. Moreover, the problems of congestion and bottlenecks are avoided because the number of clients in the community is not large. This approach is also flexible and scalable when clients are added or removed from the community. Hence, a challenging question is *how to determine such a suitable location to offer contents to the community?*

At present, many existing content distribution service providers distribute their contents by placing them on content servers which are located near the users (see *Akamai* [25]). However, the location of the contents server should not only be close to the requesting users but also be influenced by the network structures and network parameters in order to be easily found and accessed by all members of the community. Several authors like *Ouveysi et al.* [30] presented different heuristic approaches to address the video file assignment problem in VoD systems. They focused on systems with multiple file providers (herein providers are nodes that offer available video files to others) and each provider has a limited amount of local storage. *Tang et al.* [31] proposed an evolutionary approach based on genetic algorithms to solve the VoD assignment problem. These works, however, are not suitable for an application in highly dynamic and/or P2P networks, where nodes (or file providers) can be added or removed at any time. It is obvious that the approach presented in this article, i.e. moving frequently accessed files like videos in such VoD systems to super nodes in the communities, can support their quality of service requirements. The NodeRank formula as defined herein can be applied to find such suitable locations because contents can be accessed more easily from nodes with a high NodeRank that is mainly influenced by a high bandwidth of communication links. Also, it can be used in a VoD system to solve the existing accessibility problems.

B. P2P-based Distribution of Files

In P2P systems, files will be distributed among the given infrastructure, which is given by the community. When choosing a suitable location, files should be placed on the super node to facilitate their retrieval, a task for which clustering is needed. From the variety of clustering methods available, a modified ant-based approach (see Sec. II-B) will be used here, because it supports the dynamicity of large networks and works fully decentrally.

To implement the presented ideas, random walkers will travel around the network and perform the following operations:

- look for contents and files,
- pick them from low bandwidth locations

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

- and transport them to nodes on a central place with a high bandwidth and drop them there.

The distributed files will be put together on a pile of files on the super node of the community. Herein, no pheromones to direct the random walkers are used. The NodeRank values of the nodes visited by the random walkers are used for this purpose, instead. Therefore, the notion of ants is not used in the following considerations, but the notion of random walkers.

To organise the distribution of files, each node v_i in a network is assigned a NodeRank NR_i as described before and uses a limitless storage facility for files. Let F be a set of files which are located in the network. n_{F_i} is the number of files located on node i and $n_{F_{max}}$ is the maximum number of files which can be located per location. To distribute files, let A be a set of random walkers which are randomly located in the network. A random walker (with or without a file) moves from its present location v_i to a neighbour $v_j \in N_i$ selected randomly with probability $\frac{1}{N_i}$. Let $p(x)_{pick_i}$ and $p(x)_{drop_i}$ be probability functions for a random walker to pick up and to deposit a file on a node v_i.

1) Selecting Probability Functions for Picking and Depositing: In this subsection, the functions for $p(x)_{pick_i}$ and $p(x)_{drop_i}$ are considered, which are influenced by NR_i and n_{F_i} to account for the accessibility of often requested files. Three possible situations can be distinguished:

1) The node has not many files and its accessibility is poor (values of n_{F_i} and NR_i are low): It is not suitable to place a file on this node. On the other hand, it is suitable to pick up a file from this node.
2) The node has many files and is easily accessible (values of n_{F_i} and NR_i are high): This is a suitable location to deposit files, but it should be unlikely that files are picked up from here.
3) Otherwise: It is suitable to place a file on this node and pick up a file from there, depending on the value of x.

Both the number of files and the network parameters determine whether a node is a suitable candidate to drop a file there or to pick up a file from that location. Consequently, a combination x of both parameters can be defined by

$$x = \alpha n_{F_i} + \beta NR_i, \tag{8}$$

where n_{F_i} is the number of files on node i and NR_i is the NodeRank of node i. In addition, α and β are tunable parameters. Due to $n_{F_i} \in \mathbb{N}$ and NR_i in $[0, 1]$, it follows that $0 < \alpha < 1$ and $\beta \gg 1$. The value for x is strongly influenced by NR_i and n_{F_i}. If both NR_i and n_{F_i} have high values, then x will also be high and vice versa.

The functions used to determine the probabilities to pick or to drop files should behave continuously and smoothly based on the value of x. Naturally, they should return values between 0 and 1, but should never reach these values. This is a necessary requirement, because even if the dropping probability on a given node is at a high level, there will still be a tiny chance that a random walker will pick a file from there because there is always a chance that a local maximum in x can be overcome to find a better location for the files. A

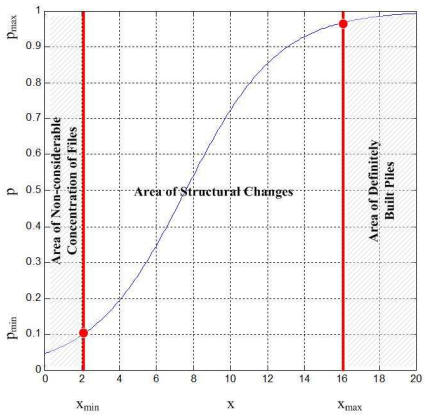


Fig. 4. Three-situation requirements of depositing probability functions

sigmoid function can therefore be applied as depositing and picking probability functions, for instance *Yang et al.* [26] deployed the sigmoid function with one adjusted parameter to define a conversion between depositing and picking by random walkers. The increase of the depositing probability is strongest for small initial values of x and saturates for large values of x. The characteristics of the sigmoid produce an S-shape, which fulfils the requirements of both probability function [27]. Linear functions for instance could only fulfill the above mentioned requirements, when they would be combined with each other. Therefore, the usage of a sigmoid function is a proper solution.

The dropping probability function is shown in Fig. 4, whereas the picking probability function simply returns the probability for the complementary event. The curve is divided into three parts: 1) initial part, where $0 \leq x < x_{min}$, 2) active part, where $x_{min} \leq x \leq x_{max}$ and 3) saturation part where $x > x_{max}$. This article consideres mainly the active part, where files will be both picked up and dropped, i.e. where structural changes take part.

According to Fig. 4, the depositing probability function is represented by the sigmoid function with two adjustable parameters, which is described by

$$p(x)_{drop} = \frac{1}{1 + e^{-a(x-c)}}, \tag{9}$$

and the picking probability function, which is

$$p(x)_{pick} = 1 - \frac{1}{1 + e^{-a(x-c)}}, \tag{10}$$

where a and c are tunable parameters.

Finally, an algorithm to pile files on the suitable place is developped.

2) Calculation of Parameters: Herein, a critical value of x is considered from the mean value of NR and $n_{F_{max}}$, which is

$$x_c = \alpha \frac{n_{F_{max}}}{2} + \beta \overline{NR}, \tag{11}$$

where $n_{F_{max}}$ is the maximum number of files which can be stored per location, and $\overline{NR}$ is the approximate average NodeRank value in the network (see Sec. III-E2).

Then α and β are calculated as follows

$$\alpha = \frac{2x_c - 2\beta\overline{NR}}{n_{F_{max}}}, \tag{12}$$

and

$$\beta = \frac{x_c}{\overline{NR}} - \alpha\frac{n_{F_{max}}}{2\overline{NR}}. \tag{13}$$

To calculate the parameters a and c, the depositing probability (Eq. 9) is considered:

$$a = -\frac{\ln[\frac{(1-p(x)_{drop_{max}})p(x)_{drop_{min}}}{(1-p(x)_{drop_{min}})p(x)_{drop_{max}}}]}{x_{max} - x_{min}}, \tag{14}$$

and

$$c = \frac{1}{2}[(x_{max} + x_{min}) + \frac{1}{a}\ln[(1 - p(x)_{drop_{max}})\frac{1 - p(x)_{drop_{min}}}{p(x)_{drop_{max}}p(x)_{drop_{min}}}]], \tag{15}$$

where $p(x)_{drop_{max}}$ is the depositing probability value for the maximum value of x, x_{max}, indicating that the node contains a pile of files and is easily accessible. $p(x)_{drop_{min}}$ is the depositing probability value for the minimum value of x, x_{min}, indicating that there are not many files here and the node's accessibility is poor.

C. Performance Evaluation

To prove the efficiency of the proposed NodeRank calculation in addressing the file distribution problem, an empirical simulation was conducted to confirm the assumption. Herein, a network with different bandwidth links was considered. A toroidal grid overlay network was utilized because of the symmetric connection of nodes. Contents (stored in files) were placed on the nodes in the network. Random walkers made a decision to pick up or place a file by considering both the current number of files and the NodeRank of the currently visited node using the formulas presented above. The aim was to place files on a node with a high NodeRank. Using NodeRank calculations, it was possible to find a suitable location for such a pile, which was easily found and accessible by the community members.

1) Simulation Results: For the simulation, the link bandwidth in a toroidal grid with 20×20 was considered. The average PageRank of this network was ≈ 0.0025. Initially, twenty files and five random walkers were placed randomly in the network. The maximum number of files that could be placed on a node was twenty.

The following parameters were used: $\alpha = 0.4$ and $\beta = 2,400$. Using Eq. 14 and Eq. 15, the parameters a and c were calculated respectively according to the values presented in Fig. 5, which were $a = 0.4$ and $c = 7.6$.

This simulation considered the large area of the low-bandwidth links. The result of the NodeRank calculations is shown in Fig. 5(a). There was a small number of nodes containing high NodeRank values. At $t = 1$, Fig. 5(b) presents the initial time of the simulation with randomly placed files and random walkers in the community. Some files were placed within the low-bandwidth links area where nodes were given a low NodeRank. Files will be placed on the suitable location based on Eq. 9 and Eq. 10. From Fig. 5(c), there are two piles of files occurring on different nodes of the community. Until $t = 13,175$, the pile of files was moved to the super node of the community that has a high NodeRank. The result is shown in Fig. 5(d).

This simulation results show that the NodeRank calculations could be applied not only to support the search but also the distribution of files. A suitable location for files can be found and selected depending on changing environmental conditions.

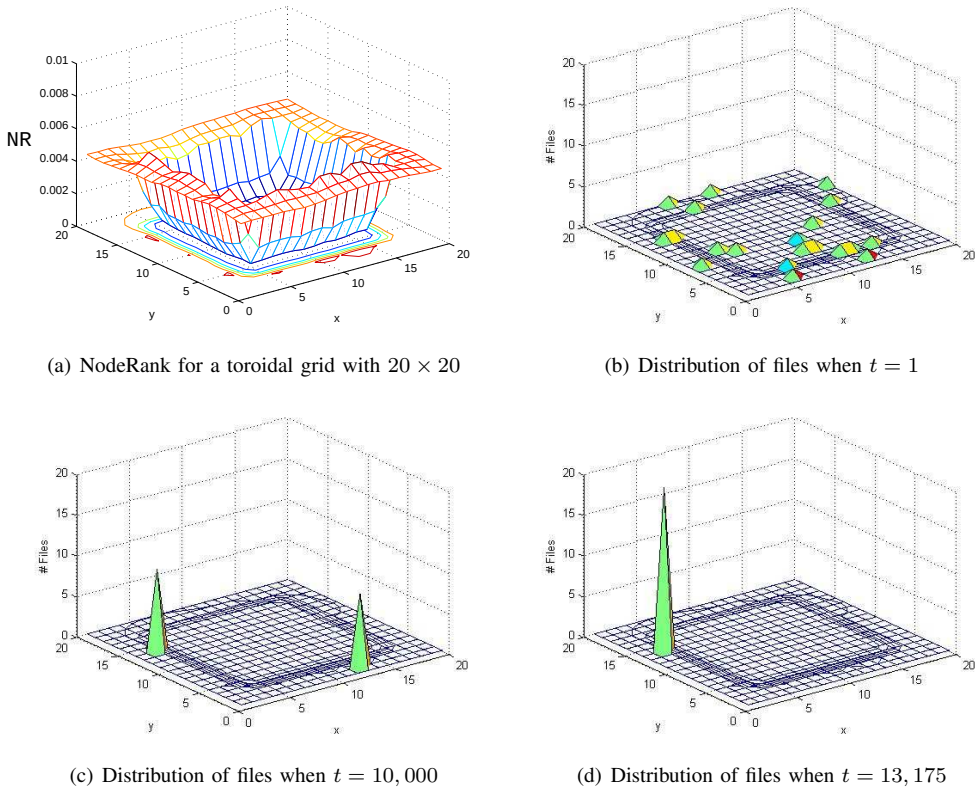
V. Conclusion

Herein, an extended PageRank calculation, which is called NodeRank, has been presented. The importance of a node is not only calculated by its position in the network graph but also by considering its network parameters. In addition, the NodeRank will be computed in a local manner using a set of random walkers. The soundness and practicability of the proposed new ideas have been evaluated by a set of simulations and their applicability in video-on-demand systems has been shown.

Nevertheless, user activity, one main factor in an information system besides network parameters and contents, will be subject for ongoing research. It is necessary to propagate user activities within the local neighbourhood and include it into the NodeRank calculation.

References

[1] L. Page, S. Brin, R. Motwani and T. Winograd, The pagerank citation ranking: bringing order to the web, *Technical report*, Stanford Digital Library Technologies Project, 1998.
[2] J. M. Kleinberg, Authoritative sources in a hyperlinked environment, *Proc. ACM-SIAM Symp. Discrete Algorithms*, pp. 668-677, 1998.
[3] Y. Joung, L. Yang and C. Fang, Keyword search in DHT-based peer-to-peer networks, *IEEE Journal. Selected Areas in Communications*, vol. 25, iss. 1, pp. 46-61, 2007.
[4] Y. Zhu and Y. Hu, Enhancing search performance on Gnutella-like P2P systems, *IEEE Trans. Parallel and Distributed Systems*, vol. 17, iss. 12, pp. 1482-1495, 2006.
[5] N. Bisnik and A. A. Abouzeid, Optimizing random walk search algorithms in P2P networks, *Computer Networks*, vol. 51, pp. 1499-1514, 2007.
[6] H. T. Shen, Y. F. Shu and B. Yu, Efficient semantic-based content search in P2P network, *IEEE Trans. Knowledge and Data Engineering*, vol. 16, iss. 7, pp. 813-826, 2004.
[7] Y. Zhu, S. Ye, X. Li, Distributed PageRank computation based on iterative aggregation-disaggregation methods, in Proc. *ACM Int. Conf. Information and knowledge management*, pp(s). 578-585, 2005.
[8] K. Sankaralingam, S. Sethumadhavan, J. C. Browne, Distributed pagerank for P2P systems, in Proc. *IEEE Int. Symp. High Performance Distributed Computing*, pp(s). 58-68, 2003.
[9] H. Ishii, R. Tempo, Distributed pagerank computation with link failures, in Proc. *the 2009 American Control Conf.*, pp(s).1976-1981, 2009.
[10] I. Stoica, R. Morris, D. Karger, F. Kaashoek and H. Balakrishnan, Chord: a scalable peer-to-peer lookup service for internet applications, *Proc. ACM SIGCOMM Conf.*, pp. 149-160, 2001.
[11] S. Ratnasamy, P. Francis, M. Handley, R. Karp and S. Shenker, A scalable content addressable network, *Technical Report*, Berkeley, 2000.
[12] A. Rowstron and P. Druschel, Pastry: scalable, distributed object location and routing for large-scale peer-to-peer systems, *Proc. IFIP/ACM Int. Conf. Distributed Systems Platforms (Middleware)*, pp. 329-350, 2001.
[13] KaZaA website: http://www.kazaa.com/
[14] J. Wang, P. Gu and H. Cai, An advertisement-based peer-to-peer search algorithm, *Journal. Parallel and Distributed Computing*, vol. 69, iss. 7, pp. 638-651, 2009.
[15] S. Milgram, The small world problem, *Psychology Today*, pp. 60-67, 1967.



(a) NodeRank for a toroidal grid with 20×20

(b) Distribution of files when $t = 1$

(c) Distribution of files when $t = 10,000$

(d) Distribution of files when $t = 13,175$

Fig. 5. Distribution of files in a toroidal grid

[16] E. Bonabeau, M. Dorigo and G. Theraulaz, *Swarm intelligence: from natural to artificial systems*, Santa Fe Institute in the Sciences of the Complexity, Oxford University Press, New York, Oxford, 1999.

[17] M. Dorigo, V. Maniezzo and A. Colorni, Ant system: optimization by a colony of cooperating agents, *IEEE Trans. Systems, Man, and Cybernetics-Part B*, vol. 26, iss. 1, pp. 29-41, 1996.

[18] J. L. Deneuborg, S. Goss, N. Franks, A. Sendova-Franks, C. Detrain and L. Chrétien, The dynamics of collective sorting robot-like ants and ant-like robots, *Proc. Int. Conf. Simulation of Adaptive Behaviour: From Animals to Animats*, pp. 356-365, 1991.

[19] E. D. Lumer and B. Faieta, Diversity and adaptation in populations of clustering ants, *Proc. Int. Conf. Simulation of Adaptive Behaviour: From Animals to Animats*, pp. 501-508, 1994.

[20] V. Ramos and J. J. Merelo, Self-organized stigmergic document maps: environment as a mechanism for context learning, *Proc. 1st Spanish Conf. Evolutionary and Bio-Inspried Algorithms*, pp. 284-293, 2002.

[21] P2PNetSim, *User's manual*, JNC, Ahrensburg, 2007.

[22] M. Zhong, K. Shen and J. Seiferas, The convergence-guaranteed random walk and its applications in peer-to-peer networks, *IEEE Trans. Computers*, vol. 57, iss. 5, pp. 619-633, 2008.

[23] C. Avin and B. Krishnamachari, The power of choice in random walks: an empirical study, *Proc. ACM Int. Symp. Modeling analysis and simulation of wireless and mobile systems*, pp. 219-228, 2006.

[24] S. Androutsellis-Theotokis and D. Spinellis, A survey of peer-to-peer content distribution technologies, *ACM Comput. Surv.*, vol. 36, iss. 4, pp. 335-371, 2004.

[25] Akamai website: http://www.akamai.de/

[26] Y. Yang, M. Kamel and F. Jin, Topic discovery from document using ant-based clustering combination, *Web Technologies Research and Development - APWeb 2005*, Lecture Notes in Computer Science, Springer Berlin / Heidelberg, vol. 3399, pp. 100-108, 2005.

[27] N. Leibowitza, B. Bauma, G. Endena and A. Karniel, The exponential learning equation as a function of successful trials results in sigmoid performance, *Journal of Mathematical Psychology*, vol. 54, iss. 3, pp. 338-340, 2010.

[28] D. Wu, Y. T. Hou, W. Zhu, Y. Zhang and J. M. Peha, Streaming video over the Internet: approaches and directions, *IEEE Trans. Circuits and Syatems for Video Technology*, vol. 11, no. 3, pp. 282-300, 2001.

[29] Y. Zeng and T. Strauss, Enhanced video streaming network with hybrid P2P technology, *Bell Labs Technical Journal*, vol. 13, iss. 3, pp. 45-58, 2008.

[30] I. Ouveysi, K. C. Wong, S. Chan and K. T. Ko, Video placement and dynamic routing algorithms for video-on-demand networks, *Proc. Global Teleommunications Conf.*, vol. 2, pp. 658-663, 1998.

[31] K. Tang, K. Ko, S. Chan and E. W. M. Wong, Optimal files placement in VOD system using genetic algorithm, *IEEE Trans. Industrial Electronics*, vol. 48, no. 5, pp. 891-897, 2001.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

Mapping relational database into OWL Structure with data semantic preservation

Noreddine GHERABI
Hassan 1 University, FSTS
Department of Mathematics and Computer Science
gherabi@gmail.com

Khaoula ADDAKIRI
Department of Mathematics and Computer Science,
Université Hassan 1er, FSTS, LABO LITEN Settat,
Morocco

Mohamed BAHAJ
Hassan 1 University, FSTS
Department of Mathematics and Computer Science
mohamedbahaj@gmail.com

Abstract— **this paper proposes a solution for migrating an RDB into Web semantic. The solution takes an existing RDB as input, and extracts its metadata representation (MTRDB). Based on the MTRDB, a Canonical Data Model (CDM) is generated. Finally, the structure of the classification scheme in the CDM model is converted into OWL ontology and the recordsets of database are stored in owl document. A prototype has been implemented, which migrates a RDB into OWL structure, for demonstrate the practical applicability of our approach by showing how the results of reasoning of this technique can help improve the Web systems.**

.

Keywords-component; ***RDB, RDF, OWL, Web ontology***.

I. INTRODUCTION

The use of ontologies is rapidly growing since the emergence of the Semantic Web. To date, the platform of Web ontologies available continues to increase at a phenomenal rate. The requirement for the development of the current web of documents into a semantic web requires the inclusion of large quantities of data stored in relational databases (RDB). The mapping of these quantities of data from RDB to the Resource Description Framework (RDF) has been the focus of a large body of research work in diverse domains. Therefore, it is necessary to study the difference between Semantic Web applications using relational databases and ontologies.

There is a need for an integrated method that deals with DataBase Migration from RDB to Object-Oriented DataBase (OODB)/XML/RDF/OWL in order to provide an opportunity for exploration, experimentation and representation of databases in a Web data. With the current revolution in the use of the Web as a platform for application development, XML (eXtensible Markup Language) [1] was the first interest to many e-business applications.

Different researches are investigated in RDB migrations focusing on different domains. Most existing proposals are restricted by a range of assumptions and characteristics such as the respect of the 3rd Normal Form and the integrity constraints [2].

Several approaches have been presented that directly map relational schemas to ontology languages [3]. Recently, the W3C RDB2RDF Working Group is developing a direct mapping standard that focuses on translating relational database instances to RDF [4].

Furthermore, in our knowledge, there are some existing work raises the issue of constructing semantic mappings between relational schemas and ontologies.

In both Database and Semantic Web communities, more and more researchers have been aware of the importance for constructing semantic mappings

In our approach we have developed a tool to create ontology from a relational database. It looks for some particular cases of database tables to determine which ontology component has to be created from which database component. This prototype extracts the schema of the database (MTRDB) then transforms it into a canonical data model (CDM) to facilitate the migration process, after the system generates the structure of OWL file and the data of RDB is stored in an OWL document

II. OUR METHODOLOGY FOR MAPPING

In order to achieve flexible mapping and high usability, we presented our approach into three separate phases, as depicted in figure 1. The first phase consists to understand the structure of the relational database and its meaning. After, the Metadata of the relational schema (MTRDB) is extracted with the Recordset of the database and in the phase 2 we develop a Canonical Data Model (CDM) to facilitate the reallocation of field values in a class structure. Finally, in the phase 3 we describe the mapping process for generating the structure and data of OWL document. At the end we present our prototype for mapping RDB into OWL.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

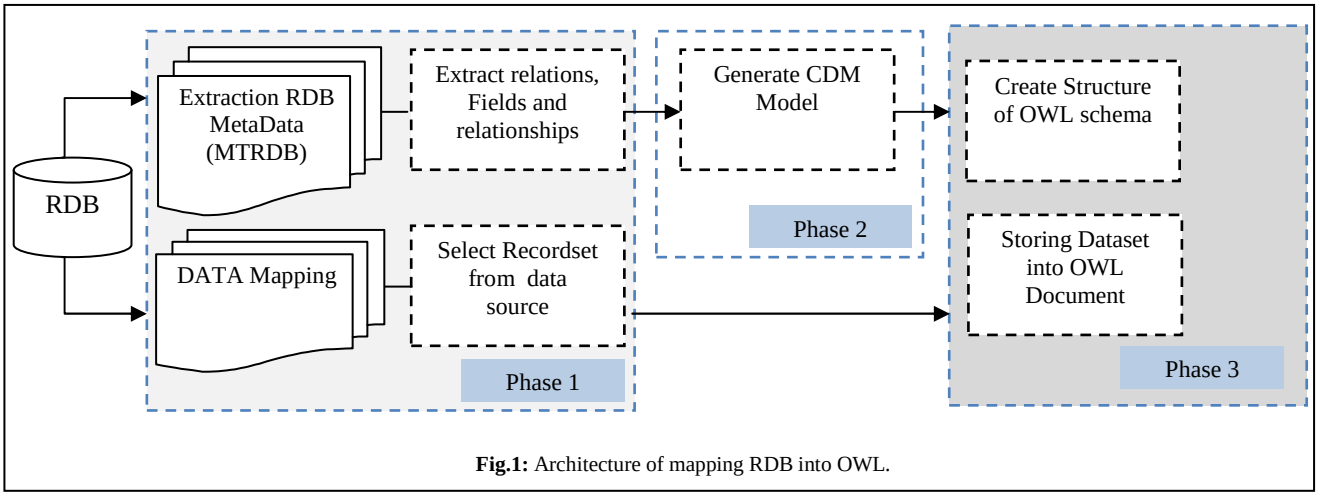


Fig.1: Architecture of mapping RDB into OWL.

A. *Mapping RDB into MTRDB*

In this section, we present the proposed process for mapping the RDB into CDM.

1) *Extracting MetaData of RDB (MTRDB).*

Our process started by extracting the basic Metadata information about the RDB, including relations and fields properties.
In our approach an RDB schema is represented as a set of elements (Relation name (R_N), set of fields (R_F), Primary Keys (R_{PK}), Foreign Keys (R_{FK}) and Unique Keys (R_{UK}))

$$MTRDB=\{R/\ R:=R_N, R_F, R_{PK}, R_{FK}, R_{UK}\}$$

- $\boldsymbol{R_N}$ is the name of the relation and R_F describes the set of fields of the relation R is defined as a set of elements:

$$R_F=\{F|F:=F_N, F_T, F_L, F_{Nl}, F_D\}$$

Where:

- **F** is the field of the relation R.
- $\mathbf{F_N}$ is the name of F.
- $\mathbf{F_T}$ its type.
- $\mathbf{F_L}$ is the data length of the field F.
- $\mathbf{F_{Nl}}$ is nullable or not.
- $\mathbf{F_D}$ denotes the default value.

- R_{PK} denotes primary key of the relation (single valued key or composite key), .
- R_{FK} denotes the set of foreign key(s) of R, $R_{FK}(R)$ = {FKn,R_{PK}(R')}, where FKn represents foreign key field name and R_{PK}(R') name of an exporting (i.e., referenced) the second relation R' that contains the referenced R_{PK}.
- Relationships (RS): A relation R has a set of relationships RS. Each relationship (rel ϵ RS) between a relation R and another relation R' is defined as:

RS(R,R') := {rel | rel := (R_{PK}(R), R, R_{FK} (R),R',Ca)}

Where R_{PK}(R) is the primary key of R, R_{FK} (R) is the foreign key representing the relationship in R' and Ca the cardinality of the source relation R

Using the DatabaseMetaData interface for retrieving the structure of the database, the table in Figure 2 shows an overview of some instructions for extraction Metadata.

MTRDB: `getMetaData`							
R: TABLE_NAME							
F_N: COLUMN_NAME				R_{PK}	R_{FK}	RS	
$\mathbf{F_T}$	$\mathbf{F_L}$	$\mathbf{F_{Nl}}$	$\mathbf{F_D}$			R_{PK}(R)	R_{FK}(R)
TYPE_NAME	COLUMN_SIZE	IS_NULLABLE	COLUMN_DEF	getPrimaryKeys	getImportedKeys	PKCOLUMN_NAME	FKCOLUMN_NAME

Fig.2: the structure of MTRDB

2) *Algorithm for extraction of MTRDB*

This section presents the algorithm for extracting *MTRDB*, is used to extract the information about MetaData of RDB, which contains the names of the relations, fields and integrity constraints of all the relations extracted from an RDB. The input to the algorithm is an existing RDB and the output is the *MTRDB* structure as described in the Section A.1. The algorithm for extraction the MTRDB from RDB is as follows:

Algorithm Extracting _MTRDB (BD: RDB) return MTRDB

MTRDB: = null; // a set to store RDB relations

For each relation r ϵ RDB do

Create element R for storing the prosperities of the relation r.

R.R_N := Extract name of (r)

For each relation R_N ϵ R do

R_N. F_N:=ExtractFieldName(R_n)

R_N. F_T:=ExtractFieldType(R_n)

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

$R_{N.}$ F_L:=Extractlengthofthefield (R_n)

$R_{N.}$ F_{Nl}:=ExtractBoolean (R_n)// (0 nullable /1 not nullable)

$R_{N.}$ F_D:=ExtractFieldDefalutValue (R_n)

End For

$R_N.R_{PK}$:=ExtractPrimaryKeys (R_n)

$R_N.R_{FK}$:= ExtractForeignKeys(R_n)

$R_N.R_U$:= ExtractUniqueKeys(r)

End For

For each set of relations (R, R') Create element RS for storing the prosperities of the relationships between R and R'.

$RS.R_{PK}$(R):= ExtractPrimaryKey (R)

RS.R:= ExtractRlation (R)

$RS.R_{FK}$ (R):= ExtractForeignKey(R')

RS.R:= ExtractRlation (R')

End For

MTRDB: = MTRDB+ R // add the relation R to MTRDB

Return MTRDB

End algorithm

B. Generating CDM from MTDATA

The next step is to define the CDM based on a classification of relations, fields and relationships, which may be performed through data access.
The CDM model is based on three concepts: class, attribute and relationship. Attributes define class structure, whereas relationships define a set of relationship types. CDM classes are connected through relationships.

CDM Class is defined as a set of classes, is denoted as 3-tuple where the first element is the name of the CDM class, the second element is a list of attributes and the latest element is the relationships between classes:

$$CDM - Class := \{C | C := (C_N, C_A, C_R)\}$$

C_N is the name of the class C, C_A is the list of attributes associated with this particular class:

$$C_A := \{A | A := (A_n, A_t, A_l, A_d)\}$$

Where A_n is an attribute name, A_t is its type, A_l is the length of this attribute and A_d is a default value if given.

C_R describes the different types of relations that can exist between any pair of classes in the CDM.

$$C_R := \{\mathrm{Re}lN, \mathrm{Re}lC, Cs, Cd)\}$$

Where $\mathrm{Re}lN$ is the name of the relationship between the source class Cs and the destination class Cd and $\mathrm{Re}lC$ is the Cardinality source of the class Cs , is represented by min..max notation.

C. OWL Structure

1) definition of OWL structure.

When the CDM has been obtained, the schema translation phase is started. Then, an appropriate set of rules is used to map the CDM constructs into OWL classes and create elements for storing OWL data
A class in OWL defines a group of individuals that belong together because they share some properties. Every individual in the OWL world is a member of the class owl:Thing. Thus each user-defined class is implicitly a subclass of owl:Thing.

Each class in CDM is translated to owl:class in the Web ontology, our class in OWL technology is represented as follows:

$$< \mathrm{owl:Class\ rdf:ID} = "\mathrm{Class} \in \mathrm{C_N}" / >$$

Each attribute A is translated into a owl:DatatypeProperty class and represented as :

$< \mathrm{owl:DataTypePr\ operty\ rdf:ID} = "\mathrm{A} \in \mathrm{C_A}" >$

$< \mathrm{rdfs:domain\ rdf:resource} = "\#\mathrm{C} \in \mathrm{C}DM\ \text{-}\ \mathrm{Class}" / >$

$< \mathrm{rdfs:range\ rdf:resource} = "\&xsd, \mathrm{Type} \in \mathrm{A_t}" / >$

$< /\mathrm{owl:DatatypePr\ operty} >$

The relationship between two classes C1 and C2, the representation of the relationship in Web ontology is represented as follows:

$< \mathrm{owl:ObjectProp\ erty\ rdf:ID} = "\mathrm{RelN} \in \mathrm{C_R}" >$

$< \mathrm{rdfs:domain\ rdf:resource} = "\#\mathrm{C_1} \in CDM\ \text{-}\ \mathrm{Class}" / >$

$< \mathrm{rdfs:range\ rdf:resource} = "\#\mathrm{C_2} \in \mathrm{C}DM\ \text{-}\ \mathrm{Class}" / >$

$< /\mathrm{owl:ObjectProp\ erty} >$

The cardinalities of a relationship are given by specifying minimum and maximum cardinalities.

For mapping the general cardinality we use:

<owl:Cardinality rdf:datatype="&xsd,nonNegativeInteger"> Cardinality ϵRelC</ owl:Cardinality>

And for mapping the maximal cardinality of each relationship we use this syntax:

<owl:maxCardinality rdf:datatype="&xsd,nonNegativeInteger"> Cardinality ϵRelC</ owl:maxCardinality>

2) Algorithm for Mapping CDM into OWL

Given a CDM Model as input, the algorithm goes through a main loop to classify CDM constructs and generate their equivalents in OWL.
The pseudo code of the mapping process is depicted in this Algorithm:

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

Input: The CDM model and Recordset of RDB

Output: The corresponding OWL schema and OWL Data

Step:

Step 1: Translate each class in the CDM model into a Class in <OWL:class>.

Step 2: Map each attribute and there proprieties in every CDM Class into <owl:DatatypeProperty> class.

Step 3: Map the relationship between CDM classes into owl:ObjectProperty class .

Step 4: Create an instance element of each recordset in RDB and translate the dataset of the recordset into instance.

Step 5: Create an OWL schema for storing CDM structure and OWL data for storing dataset**.**

EndAlgorithm

III. EXPERIMENTAL STUDY

To demonstrate the effectiveness and validity of our method, a prototype has been developed. The algorithms were implemented using Java and Oracle/Mysql.
As an example, Figure 3 shows a relational database, PKs are bold and FKs are marked by "#".

- Product(**ProductID**, ProductName, ProductPrice)
- Customer(**CustomerID**, CustomerName, CustomerAdress)
- Employee(**EmployeeID**, EmployeeName)
- Order(**OrderID**, OrderDate, OrderQuantity, #CustomerID, #ProductID, #EmployeeID)
- Store(**StoreID**,StoreName)
- EmployeeStore(#**EmployeeID**,#**StoreID**

Fig. 3. Sample Relational database

The Conversion phase consists to converting existing RDB data to the text format defined by the target schema. Data stored as tuples in an RDB are converted into complex individuals in OWL document. We propose using CDM to guide the conversion process. Firstly, the RDB relations tuples are extracted using MetaDatabase instances. Figure 4 shows the RDB structure extracted from database. Secondly, these data are transformed to match the target format. Finally, the transformed data are stored into text files.

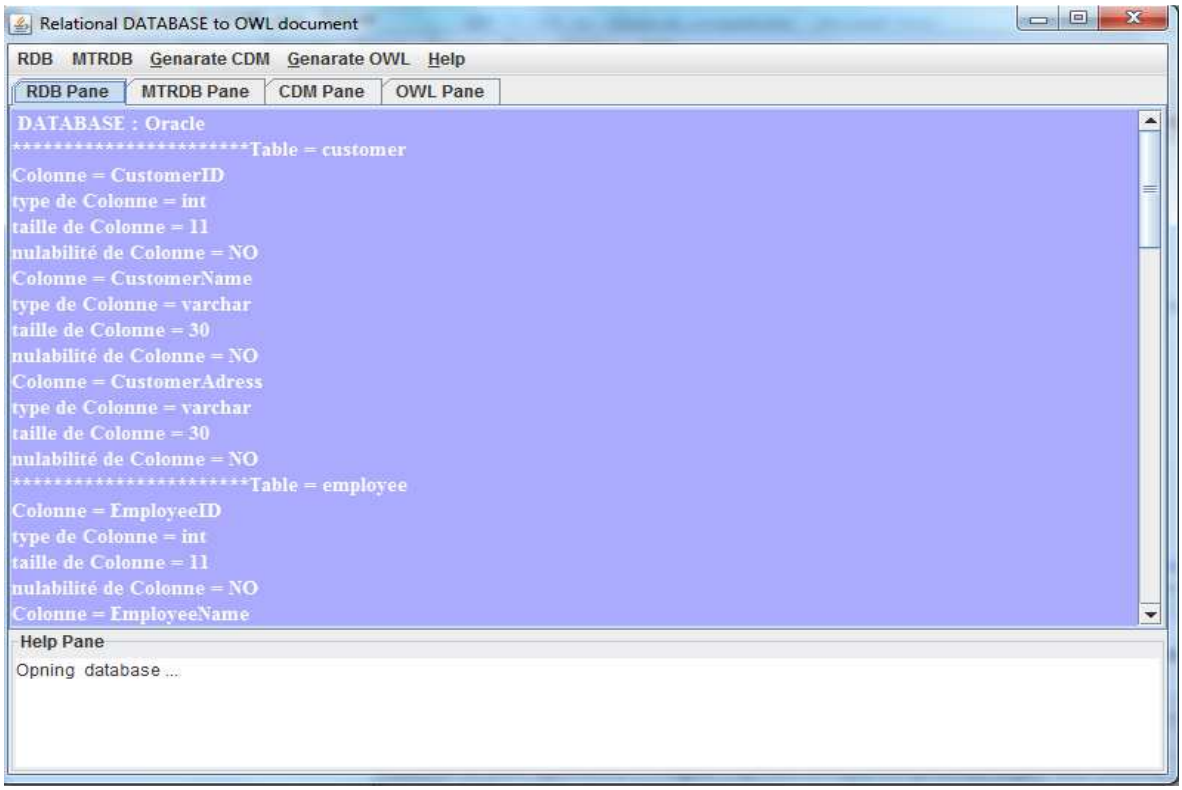


Fig.4: Extracting the RDB structure from database.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

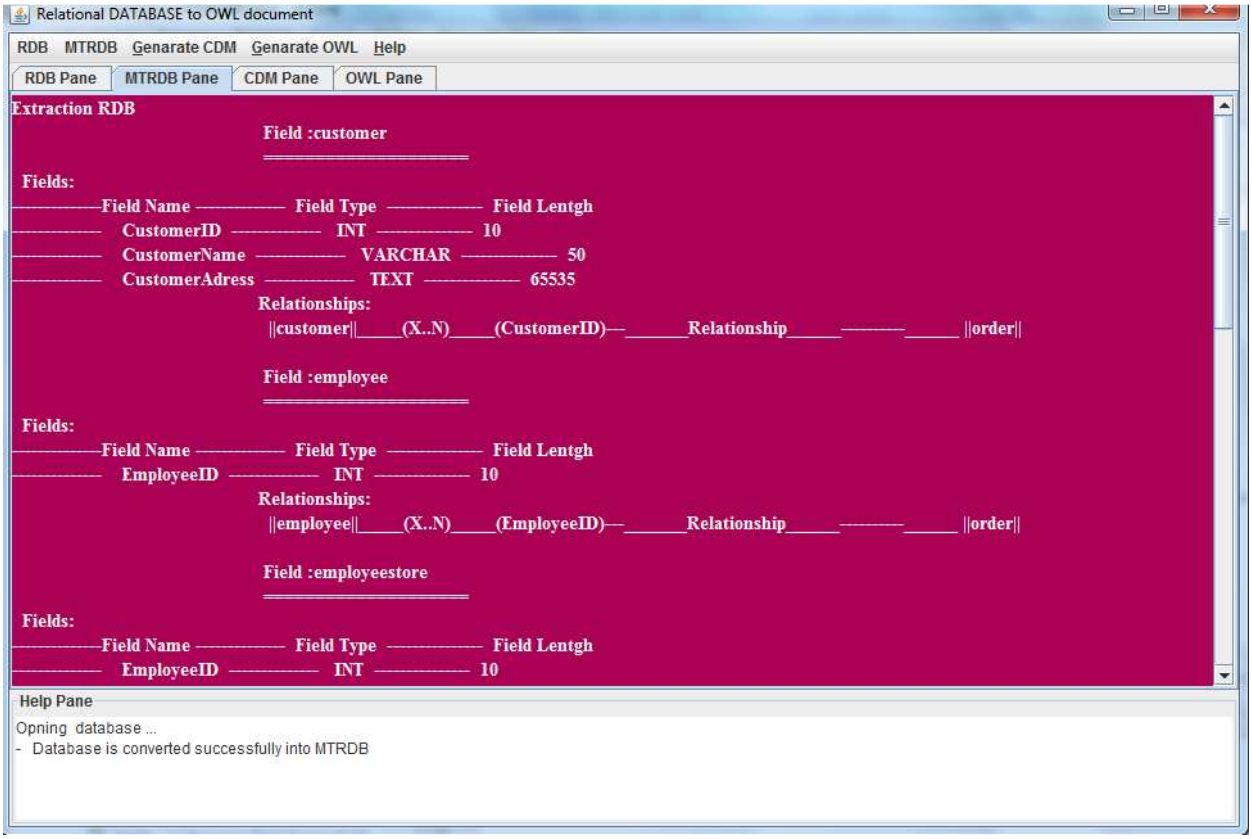


Fig.5: The MTRDB structure

The algorithm classifies each relation in the MTRDB by matching its attributes, primary key, foreign keys and its constraintes, and then maps the relation into CMD classes. Figure 5 shows the structure MTRDB extracted from RDB

During the mapping process, a CDM structure is automatically generated by the system to record the relationships between generated ontology components and the original database components, as shown in the platform of Figure 6

IV. Related work

In recent years, with the growing importance and benefits provided by Web semantic, there has been a lot of effort on migrating RDBs into the relatively newer technologies (XML/RDF/OWL) [5], [6], [7], [8]. Before applying a method for mapping relational database into web ontology, it must first extract the conceptual schema relational model. Extracting conceptual schema from a logical RDB schema has been extensively studied [9], [10]. Such conversions are usually specified by rules, which describe how to deduce RDB constructs (e.g., relations, keys), classify them, and identify the relationships. Fonkam et al [11] propose also an algorithm for converting RDB schemas into conceptual models

Blakeley [12] proposes a method for mapping RDB, this method consist to generate mappings between RDB and RDF with the RDB table as a RDF class node and the RDB column names as RDF predicates. Cullot et al [13]. use an efficient method for generating classes from tables and converts column to predicate, by using the specific relational database schema characteristics, after the mappings are stored in a R2O document.

V. Conclusion

In summary, the main achievements of this paper are listed as follows. Firstly, we have presented a new approach for mapping relational database into Web ontology. It captures semantic information contained in the structures of RDB, and eliminates incorrect mappings by validating mapping consistency. Secondly, we have proposed a new algorithm for constructing contextual mappings, respecting the rules of passage, and integrity constraints.
Finally, we have experimentally evaluated our approach on several data sets from real world domains. The results demonstrate that our approach performs well as compared to some existing approaches in average.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

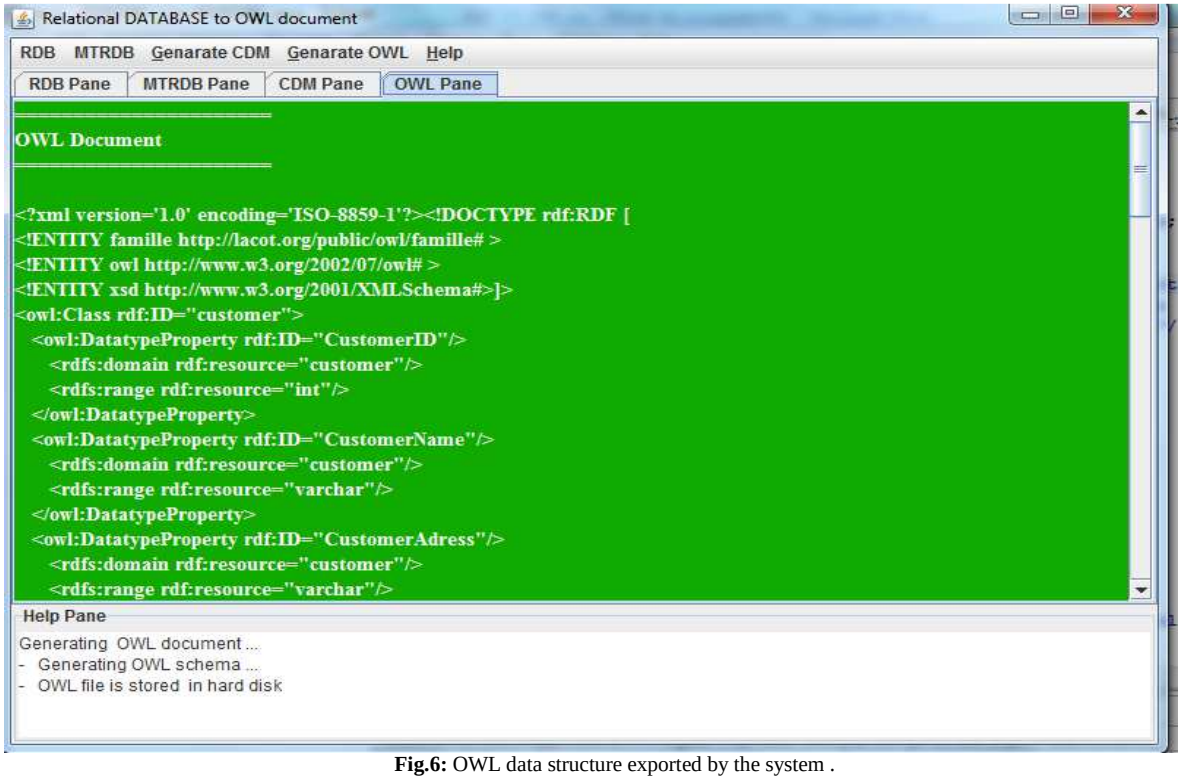


Fig.6: OWL data structure exported by the system .

VI. REFERENCES

[1] W. J. Pardi, XML in Action, Microsoft Press, Washington, 1999.

[2] Fahrner, C. and Vossen, G.: Transforming Relational Database Schemas into Object-Oriented Schemas According to ODMG-93.In 4th Int. Conf. on Deductive and Object-Oriented Databases,pp. 429–446, 1995.

[3] J. F. Sequeda, S. H. Tirmizi, O. Corcho, and D. P. Miranker. Survey of directly mapping sql databases to the semantic web. Knowledge Eng. Review, To Appear 2012

[4] M. Arenas, E. Prud'hommeaux, and J. Sequeda. Direct mapping of relational data to RDF. W3C Working Draft 24 March 2011, http://www.w3.org/TR/rdb-direct-mapping/.

[5] Green, J., Dolbear, C., Hart, G., Engelbrecht, P., Goodwin, J."Creating a semantic integration system using spatial data", , in *International Semantic Web Conference 2008* Karlsruhe, Germany

[6] Noreddine Gherabi and Mohamed Bahaj. Robust Representation for Conversion UML Class into XML Document using DOM. *International Journal of Computer Applications* 33(9):22-29, November 2011

[7] Cristian P´erez de Laborda and Stefan Conrad. Relational.OWL - A Data and Schema Representation Format Based on OWL. In Second Asia-Pacific Conference on Conceptual Modelling (APCCM2005), volume 43 of CRPIT, pages 89–96, Newcastle, Australia, 2005.

[8] Tirmizi et al, "Translating SQL Applications to the Semantic Web", Tirmizi, S., Sequeda, J., Miranker, D., Lecture Notes in Computer Science, Volume 5181/2008 Database and Expert Systems Applications-(2008)

[9] Wu, Z., Chen, H., Wang, H., Wang, Y., Mao, Y., Tang, J., Zhou, C., "Dartgrid: a Semantic Web Toolkit for Integrating Heterogeneous Relational Databases", Semantic Web Challenge at 4th International Semantic Web Conference (ISWC 2006), Athens, USA, 5-9 November 2006.

[10] Alhajj, R.: Extracting the Extended Entity-Relationship Model from a Legacy Relational Database. Inf. Syst, vol. 28, pp. 597–618, 2003.

[11] Fonkam, M. M. and Gray, W. A.: An Approach to Eliciting the Semantics of Relational Databases. In 4th Int. Conf. on Advanced Info. Syst. Eng., vol. 593, pp. 463–480, 1992.

[12] Blakeley, "RDF Views of SQL Data (Declarative SQL Schema to RDF Mapping)", Blakeley, C., OpenLink Software, 2007.

[13] Cullot, N., Ghawi, R., Yetongnon, K..,"DB2OWL: A Tool for Automatic Database to Ontology Mapping", In Proc. of 15th Italian Symposium on Advanced Database Systems (SEBD 2007), pages 491-494, Torre Canne, Italy, 17-20 June 2007.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

A Three-Layer Access Control Architecture Based on UCON for Enhancing Cloud Computing Security

Niloofar Rahnamaee
Department of Computer Engineering
Tehran North Branch, Islamic Azad University
Tehran, Iran
niloofar_rahnamaee@gmail.com

Ahmad Khademzadeh
Scientific and International Cooperation Department
Iran Telecommunication Research Center
Tehran, Iran

Ammar Dara
Department of Computer Engineering
Science and Research Branch, Islamic Azad University
Tehran,Iran
ammar.dara@gmail.com

***Abstract*— By emerging cloud computing, organizations utilize this new technology by consuming cloud services based on-demand. However, they must put their data and processes on a cloud, therefore; they do not have enough control on their data and they must map their access control policies on access control policies of a cloud service. Also, some aspects of this technology like interoperability, multi-tenancy, continuous access control are not supported by traditional approaches. The usage control model with two important specifications like continuous access control and attribute mutability are more compatible with security requirements of cloud computing. In this paper, a three layer access control based on the usage control for could services has been proposed, in which separation of duties can support the multi-tenancy and the least privilege principle.**

Keywords-Clould Computing; Access Control; Usage Control (UCON); Multi-tenancy; Separation of Duties

I. INTRODUCTION

Cloud computing, as an innovational improvement in IT technology, is a revolution in the software industry [1]. The main goal of cloud technology is to realize "network as a high performance computer" [2] in a way that all users, are capable to running processes and storing data on this infrastructure. Instead of traditional approaches, on-demand services will deliver with a lower cost for organizations [2]. To achieve this, all data and processes should move onto cloud, which normally results in less security controls of the organization on its own data and processes. However, organizations prefer to access to their own data and processes with their own policies[1]. According to openness, distribution and non-heterogeneity [1][2][3] nature of cloud, data integrity, confidentiality, privacy[3] and authorization[4] may be in danger. Access control as a security mechanism guarantees that a specific resource just and only is accessed by an authorized user [5].

Many different access control schemes has been offered for distributed systems, but the attribute-based models look more appropriate [1][2][3][5][6][9].

There are three requirements for cloud services as follows:

1) Cloud service must be able to specify access control policies of end users to service objects, which is based on its business logic.

2) A cloud service consumer must be able to enforce more access control policies on its user requests to the objects of the organization. When an organization wants to use a cloud service, it must map its policies on access control policies of the cloud service. This mapping of policies may violate the least privilege principle. Therefore, organization can prevent violating their policies by enforcing more policies on access requests.

3) Cloud service vendor must be able to offer cloud services to consumer in all applicable levels. For example, tenants may rent only necessary functions with a lower cost instead of all the services.

According these three requirements, the usage control model is the best option among various access control policies. In this paper, a three level architecture based on the usage control model is presented, which not only uses separation of duties but also supports multi-tenancy and cross-domain communication.

In the second section of this paper, previous works and researches on this subject are considered. Then, proposed approach based on a three-layer access control is explained in the third section. Section 4 describes the architecture of a three-layer access control model based on usage control along with four components. Then the proposed architecture has been analyzed. We will give a conclusion description finally in section 5.

II. RESEARCH AND RELATED WORKS

According to the nature of cloud computing which is extensible, heterogeneous and multi-tenant, it is necessary to consider these specifications in access control policies.

Xiao and associates use access control list (ACL) to support multi tenancy [1]. In this research, access control is divided into different levels: cloud service provider and tenant. The service provider creates a record per each tenant so that include an managerial <s,o,a> tuple which tenants can manage their users, objects and ACL by means of it. Jose M. Carlo and associates offered an especial authorization model for cloud computing which customized the access control on a federated environment for organization cooperation [4]. In this model

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

the authorization 3-tuple <Subject, Privilege, Object> expanded to 5-tuple < Issuer, [User|Role], privilege, interface, ObjectPath > which are explained as follows: Issuer defines that the User | Role has sufficient privilege on ObjectPath via interface.
Also for assigning the membership of user to a role and supporting hRBAC, the tripe <Issuer,[User|Role], roleName> has been defined which explained as: Issuer defines that the User | Role is responsible for the role/sub-role with role name. Therefore, the organizations define their access control policies to their own resources on cloud, by these 5 and 3 tuples [4].

Chen Danwei and associates offered access control architecture according to usage control model (UCON) for cloud computing. The majority of this paper is a negotiation module in authorization architecture to improve the flexibility of access control on cloud services. When the requested access has not sufficient attributes, a second access choice via a negotiation module will provide, rather than of refusing access directly [2].

The general scenario of access control UCON model defined by three specifications in Fig.2. This scenario, divides the usage control in three phases: before usage, ongoing usage, and after usage. Decision-making control components (Authorization, Obligations and Conditions) can check and enforce in first two phases [7][8][9][10]. Obligations will not consider on after usage phase in Core UCON Model, but in papers [11][12] post-obligation are extended for Core UCON Model. In this paper we use the extended model of UCON.

UCON is a session based access control model, because it controls not only access request, but also the ongoing access. Mutability means that the attributes of objects or subjects can be updated as a result of an access. There are three types of updates: pre-update, on-update and post-update. Updating the attribute of an object or subject may result in to allow or revoke current access or another access, according to the authorizations of the access [13].

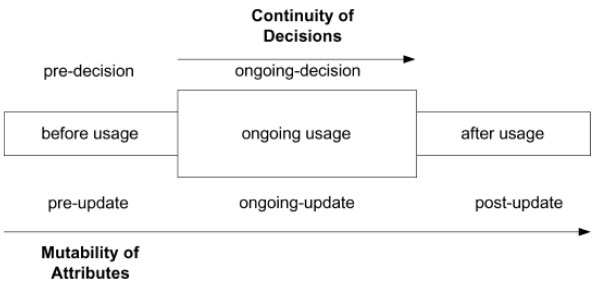


Figure 1. UCON scenario [13]

There are three main actors in cloud environment: user, vendor and original cloud provider which will consider as tenant, service vendor and service creator, respectively. Tenant is an organization that rent the cloud from cloud service vendors and it can have users.

The cloud vendor is an organization that offers the cloud services to the cloud user with guaranteed quality of experience (QoE) and quality of service (QoS) within the framework of a service level agreement (SLA). Service creator is a developer service organization which provides access for tenants' users to its services via service vendors.

III. THE PROPOSED THREE-LAYER ACCESS CONTROL

In the cloud environment, service creators usually define access control policies of end users to a cloud service. However, tenants usually tend to have the most possible control on their data and be able to enforce more policies than by the service creators on their access request of their end users. In addition, vendors tend to offer their services to consumers in all desired levels. Therefore, cloud access control mechanism must be able to support these three requirements. As a result, in this paper; a three-layer architecture is proposed for decision and enforcement of access control policies. the layers are as follow:

- *Service layer*: as an enforcer of service access control policies.
- *Provider layer*: as an enforcer of vendor access control policies.
- *Tenant layer*: as an enforcer of service consumer access control policies.

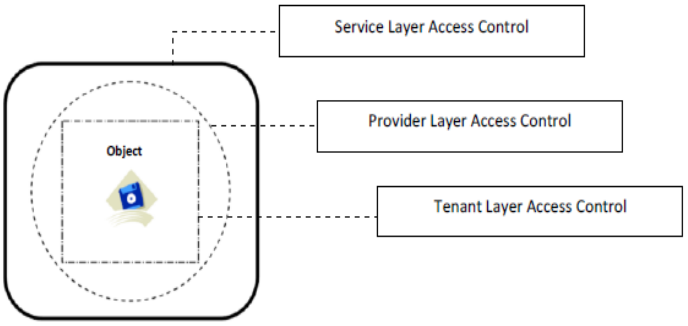


Figure 2. The enforcement of three layers access control on user's objects

In the service layer, it has been guaranteed that service objects will be available for end users, according to creator access policies. The service creator specifies these policies based on a business logic, which is related to that service. For example, in a healthcare service, the service creator will determine rights of the *doctor* role. In this layer, creators assign the first limits of the access rights of a cloud service. Therefore, these policies specify the maximum rights of other layers.

In the provider layer, a service vendor can offer its service to its tenants in various levels. Some of service usage contracts are enforced by access control policies in this layer. Vendors define access rights of their tenants. For example, hospital A can rent a healthcare service only for its laboratory, while hospital B not only want to rent the laboratory, but also for prescription and diagnosis sections. Therefore, further than creator layer policy limitations, more policy enforcement is possible. Hence, more limitations are enforced than service layer.

In the tenant layer, organizations can enforce more policies to the previous layers. Then the least privilege principle can be applied for all their users and objects in a

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

cloud. Usually, organizations using a cloud service have their own interior access control policies. Therefore, they must map their policies on cloud service policies. This mapping may violate the least privilege principle for some users. It may permit an unauthorized access according tenant policies; although it is permitted cloud service policies. Hence, tenants can enforce more policies than policies by a cloud service creator and vendor. For example, service creator of a healthcare service permits billing right for *nurses*, however hospital A does not want their nurses have this right. Therefore, Hospital A must map the nurse role to the nurse role of the service with billing right, which violates the organization polices and the minimum privilege principle. Hence, hospital A tends to have a nurse role without billing rights. Hospital A can revoke nurses' billing rights in the tenant layer. In the tenant layer, more limits can be enforced other than two previous layers.

As shown in Fig. 2, preliminary access rights are defined in the first layer. Then, vendor layer can restrict the service layer of access rights, and finally tenant layer can limit access rights of the two previous layers.

IV. THREE-LAYER ARCHITECTURE BASED ON UCON FOR ACCESS CONTROL OF A CLOUD SERVICE

In this paper, a three-layer access control architecture based on the usage control is proposed, which is "platform as a service" and guaranties access control for SaaS services.

In the Fig. 3, four components of the proposed access control architecture are shown, which are as follow:

- Access control service
- Service provider
- Cloud provider
- Identity provider.

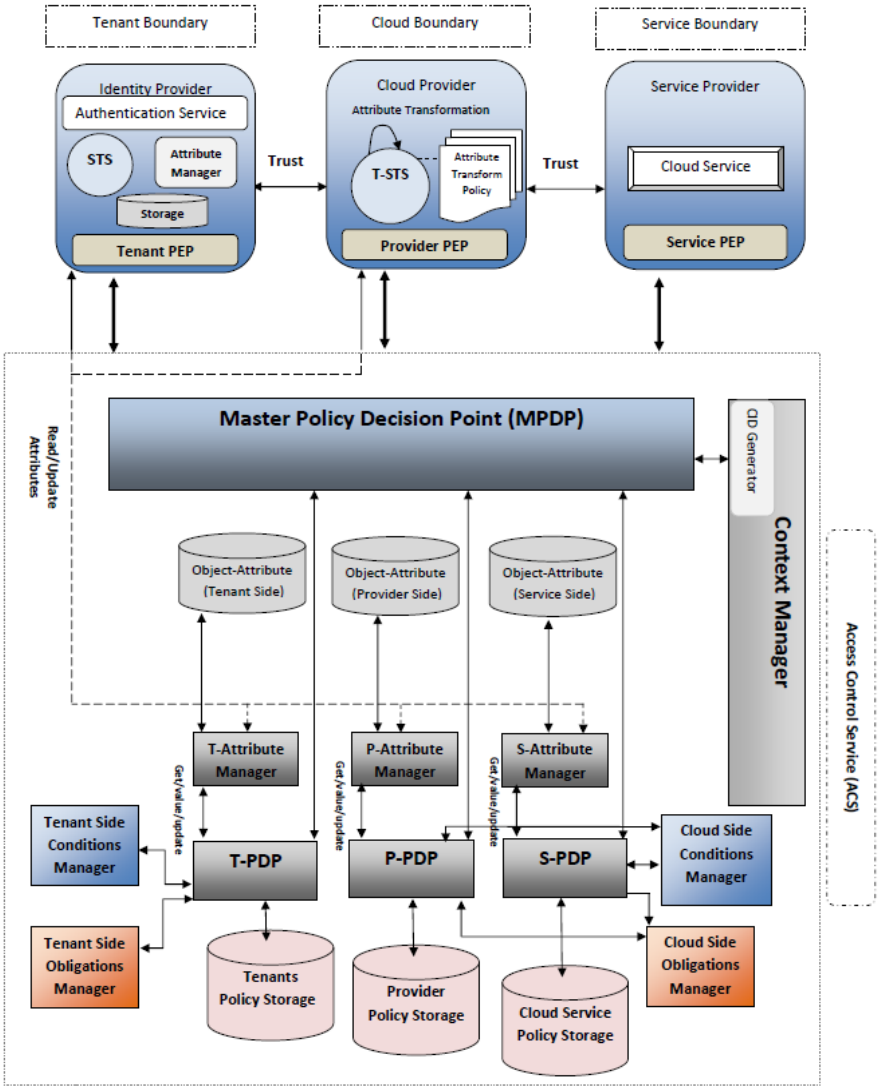


Figure 3.The architecture of proposed three-layer usage control in cloud

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

A. *Access Control Services (ACS)*

Access control service as a decision engine is the major component of this architecture (see Fig. 4). This service, receives access requests to a cloud and decides about their permit access.

This component consists of a master policy decision point (MPDP) and three other deciding components as Tenant PDP (T-PDP), Provider PDP (P-PDP) and service PDP (S-PDP) respectively for tenant, provider, and service levels. Each PDP has its own attribute manager and policy storages for its specific layer. There are three policy storages: object attributes of tenant, consumer and provider.

Another component along with MPDP is the context manager, which creates a record of an access control state for any request.

Tenant and cloud side of condition managers are responsible for managing the conditions of tenant and cloud, respectively. Similarly, Tenant and cloud obligation managers are considered for managing the obligations of tenants and clouds.

The MPDP component decides to invoke which lower level PDP for deciding, after receiving a request. For any request, the MPDP may only invoke a PDP or all three PDPs.

The T-PDP retrieves the proper policy storage after receiving a request from the MPDP. It receives attributes, conditions and obligations from tenant side components in ACS. After the verification of authorization, the results is sent back to MPDP. The S-PDP and P-PDP operates like the T-PDP.

In the ACS, three components are considered for managing the attributes, which are respectively for updating and retrieving of T-attribute, V-attribute, and S-attribute managers of tenants, vendors and service providers.

B. *Identity Provider*

Identity provider is a component that tenants are connected to ACS using it. Requests of users for consuming a cloud service first are sent to this component. After passing authentication verification step, security token service (STS) creates a token and sends it to ACS through tenant PEP. Tenant PEP is responsible for enforcing the tenant policies on the requests.

C. *Cloud Provider*

Another component of this architecture is cloud provider. The Transform Security Token Service (T-STS) component in the cloud provider, not only has trust with STSs of tenants and cloud services, but also is responsible for translating of inter-organizations' tokens. By interfering of this component between two components of token and service providers, interoperability has been better through mapping of attribute transformation. In fact, the T-STS have a set of token translation policies, which can perform this mapping. An access request along with token are sent to ACS through the PEP provider. The PEP provider is responsible for enforcing vendor policies.

D. *Service Provide*

The service provider component is responsible to enforce service creator policies by means of ACS. Therefore, this component is not engaged with vendor and tenant policies. The Service PEP is responsible for enforcing service policies on access requests.

E. *Access control steps in the proposed architecture*

Communication between components is shown in Fig. 4. First in the tenant layer, end users send their request for using an object to identity provider of their domain (step 1). The identity provider using its STS issue token and send it to the MPDP along with the access request through tenant PEP. Then, the MPDP invoke T-PDP for verifying tenant layer access control (step 2). After that, the T-PDP retrieves the necessary policies from tenant policy storage and using attribute manager, provides necessary attributes for authorization of the access control. Then, it verifies obligations and conditions using obligation tenant side condition and obligation managers. If tenant layer policy permits the access request of the user, it sends back access permit to the tenant PEP. Otherwise, it sends back access denied (step 3). If the identity provider receives the *access permit* message, it sends the access request along with the user token to the cloud provider (step 4). If it receives *access deny* message, it skips the execution of the access. Therefore, inter organization access of users to a cloud service are determined in this layer.

By receiving the access request and its token, the cloud provider translates attributes in the token to proper attributes for provider layer using T-STS and its attribute transform policy and creates a new token. Then, the provider PEP sends the access request and its token to the MPDP in ACS, the MPDP invoke the P-PDP for verifying the access request of the provider layer (step 5).

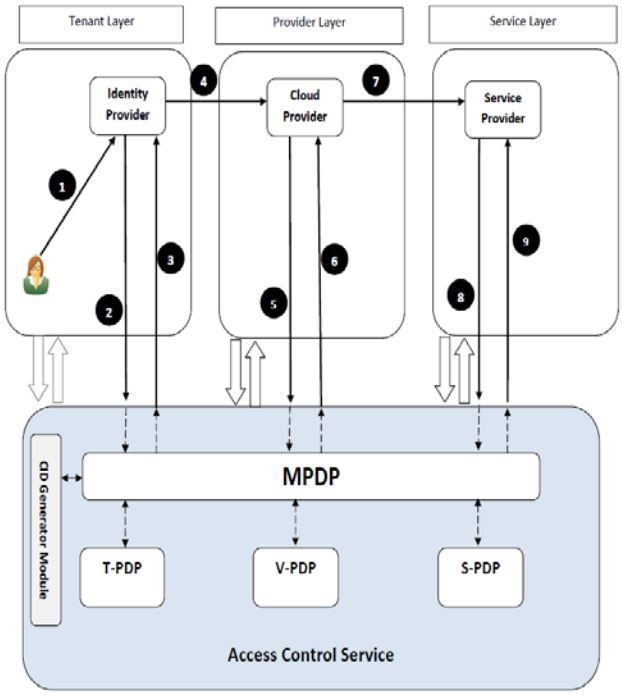


Figure 4. Access control steps in the proposed architecture

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

Therefore, the P-PDP retrieves the necessary policy from provider policy storage and using P-attribute manager, provides needed attributes for authorization of the access request. It verifies the obligations and condition using cloud side condition and obligation managers. If the policy of the provider layer permits the access request, P-PDP sends the access permit to the MPDP. Then, the MPDP sends it to the provider PEP, Otherwise, access denied is sent (step 6). If the cloud provider receives the access permit, it translates the attributes to proper attributes for the service layer using attribute transform policy. Afterwards, it sends the access request and it token to the service provider component; otherwise, it sends *access deny* message for skipping the execution of the access. In this layer, the tenant access request to a cloud service is controlled by vendor policies.

By receiving the access request and its token, the cloud provider sends it to the MPDP in ACS using the service PEP. The MPDP invokes the S-PDP for verifying the access request of the service layer (step 8). Afterwards, the S-PDP retrieves the necessary policy from the service policy storage. Using S-attribute manager, it provides the necessary attributes for authorization of the access request. Using cloud side condition and obligation managers, it considers the conditions and obligations. If the policy of the service layer permits the access request, the S-PDP sends an access permit to the MPDP. The MPDP sends it back to the service PEP, otherwise; it sends access denied (step 9). If the service provider receives the access permit, it allows the execution of the access request, if not; it skips it. In this layer, an access request of a user to an object is controlled by a service policy.

V. ANALYSIS OF THE PROPOSED METHOD

This three-step architecture has some advantages as follows: First, a service provider may put its service on cloud based on its own policies without any worries that what organizations with what policies may use it. In addition, there are no difficulties about translation of attributes for different layers. This specification encourages the service providers to put their services on the clouds without any worry about their interoperability.

Second, tenants may enforce more policies than policies by cloud services on their users. Therefore, they may define more policies on the cloud service policies for respecting the least privilege principle. It means the maximum control of a tenant on using a service in a cloud. Hence, the tenants are not worry about mapping their access control policies. In addition, there is no difficulty about translating of attributes for cloud services and their interoperability.

Third, considering the provider layer can allow the vendors to enforce their policies according to their agreements other than the policies defined by the cloud service policies. Therefore, they can a enforce security policies in various levels. This means that there is the maximum flexibility for the vendors for using a cloud service. Also, there is no difficulty about the translation of the attributes, the extendibility and scalability of tenants in the cloud. Because there is no conflict regarding inhomogeneous of policies and attributes in the different domains, thus there is a better interoperability.

Finally yet importantly, because the access control policy in ACS is based on the usage control model, two built-in and important specification of this model continuous access control and mutability exist. Therefore, a vast range of policies may be defined. Also, the usage control is performed during access request, furthermore, it is checked during an ongoing access. If during an access, the policies violate, the access can be revoked from the user.

VI. CONCLUSION

In this paper, a three-layer access control based on a usage control model for cloud services has been proposed. This architecture, for considering the least privilege principle, increasing of cross domain interoperability, data control and process of tenants by themselves has been presented. In addition, vendors can offer their services in various levels for a specific cloud service.

I. REFERENCES

[1] X. Li, Y. Shi, Y.Guo, and W. Ma, “Multi-Tenancy Based Access Control In Cloud,” IEEE Conference, 2010.

[2] Ch. Danwei, H. Xiuli, and R. Xunyi,”Access Control of Cloud Service Based on UCON, Cloud Computing,” Springer,Berlin, 2009.

[3] Kh. M. Khan, and Qutaibah Malluhi, “Establishing Trust in Cloud Computing,” IEEE IT Pro Journal, 2010.

[4] M. Joes and Others, “Toward a Multi-Tenancy Authorization System for Cloud Services,” Computer and reliability society IEEE, 2010.

[5] P. Samarati1 , and S. d. C. di Vimercati, “Access Control Policies, Models, and Mechanisms,” FOSAD Springer ,2001.

[6] A. Dara, F. Shams, P. Mehregan, “An Access Control Model Based On Language Theory For Service Oriented Architecture,” International conference communication and information security IASTED,2010.

[7] J. Park, and R. Sandhu, “The UCON-ABC Usage Control Model” , ACM transaction on Information and System Security,2004.

[8] A. Lazouski, F. Martinelli, and P. Mori, “Usage control in computer security - A survey,” Elsevier, 2010.

[9] M. Menzel , C. Wolter, and C. Meinel, “Access Control for Cross-Organisational Web Service Composition,” Journal of Information Assurance and Security, 2007.

[10] M. Colombo, A. Lazouski , F. Martinelli, and P. Mori, “A Proposal on Enhancing XACML with Continuous Usage Control Features,” Springer, 2010.

[11] A. K. Talukder, and L. Zimmerman, “Cloud Economics- Principles, Costs, and Benefits,”Springer, 2010

[12] M. Colombo, A. Lazouski, F. Martinelli, and P. Mori, “A Proposal on Enhancing XACML with Continuous Usage Control Features,”Springer, 2010.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

Detection of DoS and DDoS Attacks in Information Communication Networks with Discrete Wavelet Analysis

OlegI. Sheluhin
Department of Information Security
Moscow Tech. Univ. of Communication and Informatics
Moscow, Russia

Aderemi A. Atayero
Department of Electrical and Information Engineering
Covenant University
Ota, Nigeria

***Abstract*—A method based on discrete wavelet decomposition of traffic data and statistical processing algorithms based on Fisher and Cochran criteria are proposed for detection of traffic anomaly in computer and telecommunication networks. Two sliding windows with two different threshold values are employed to reduce the level of false alerts. A high efficiency level of detection of abnormal traffic spikes is thus guaranteed. The paper likewise presents an algorithm developed for detecting DoS and DDoS attacks based on these statistical criteria. Software is developed in *Matlab* based on the proposed algorithm. Data sets made available by the Lincoln Laboratory of MIT (1999 DARPA Intrusion Detection Evaluation) were analyzed as the test sequence. Analysis of experimental results revealed that the ultimate test for detecting an attack is to check if any one of the statistical criteria exceeds the upper threshold at the stage of coefficients reconstruction.**

Keywords-Anomaly, Denial of Service, DDoS, Wavelet transform, DWT, FWT

I. INTRODUCTION

Statistical methods for detecting network attacks are based on a comparison of the statistical characteristics of packet flow, averaged over a relatively short period of time (local characteristics), with appropriate characteristics for an extended period of time (global data) [1 - 4]. If the local characteristics differ significantly from the corresponding global characteristics, it is indicative of an anomalous behavior of packet flow, and an attempt to scan the network or network attack is highly probable. The problem thus arises of constructing effective methods for calculating the local statistical characteristics for a limited period of time and determination of local characteristics of the anomalous deviation from the global statistical characteristics of the packet flow.

We propose in this paper a method for solving the problems of traffic anomaly detection in computer and telecommunication networks based on discrete wavelet decomposition of traffic data and statistical detection algorithm using Fisher's and Cochran criteria [5]. The article also examines the harbingers of abnormal packet flow in the network and the relationship between these harbingers using different statistical criteria.

Datasets provided by the Lincoln Laboratory Massachusetts Institute of Technology (1999 DARPA Intrusion Detection Evaluation) were obtained and used in the analysis, representing the network traffic collected at the border router of the university network [6]. Each sequence spanning approximately 24 hours with discretization step of 1s is presented as pure 'unadulterated' network traffic without attack, as well as in the form of adulterated traffic with different types of anomalies relating to attacks such as denial of service (DoS) and different types of unauthorized network sniffing. DoS attacks also incorporate distributed DoS attacks (DDoS), which entail the 'owning' of a number of unsuspecting host computers for the purpose of stealthy attacking a targeted single victim computer [7].

II. DISCRETE WAVELET TRANSFORM: MALLAT ALGORITHM

Huge costs in computational power will be incurred for calculating the wavelet spectrum with continuous change of the s and u parameters. The set of $\psi_{us}(t)$function has a high level of redundancy. Discretization of these parameters becomes necessary with the possibility of restoring a signal from its transformation. Discretization is usually carried out in powers of two as given in (1):

$$\psi_{j,k}(t) = \frac{1}{\sqrt{s}}\psi\left(\frac{t-u}{s}\right) = \frac{1}{\sqrt{2^j}}\psi(2^{-j}t - k) \qquad (1)$$

where $s = 2^j$, $u = k2^j$, j and k – whole numbers.

In this case, the *u, s* plane is into the corresponding *j,k*grid. The parameter *j* is *the scale parameter or the level of decomposition;* the wavelet transform performed with such scale parameter is called *dyadic.*The fastest and most commonly used discrete wavelet transform is the so-called fast wavelet transform (FWT) or Mallat algorithm [8].In accordance with the Mallat algorithm, a signal can be represented as a set of successive rough approximations $A_j(t)$ and exact *(detailed)* $D_j(t)$ componentswith their subsequent refinement using the iterative method (2).

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

$$S(t) = A_j(t) + \sum_{j=1}^{m} D_j(t) \quad (2)$$

Each refinement step corresponds to a given scale 2^j (i.e. index *j)* of analysis (decomposition) and synthesis (reconstruction) of the signal.Such wavelet representation of each component of the signal can be viewed both in the time and frequency domains. For example in the first step of the algorithm, the input signal *S*(*t*) decomposes into two components (3):

$$S(t) = A_1(t) + D_1(t) = \sum_k \alpha_1 \phi_1(t) + \sum_k d_1 \psi_1(t) \quad (3)$$

where $\psi_{1k}(t)$ - wavelet, $\phi_{1k}(t)$ - wavelet generating function, a_1,d_1– Coefficients of the *approximate* and *detailed* components at level 1, respectively.

One of the advantages of wavelet transform is that it provides an opportunity to analyze the signal in the frequency-time domain, thus allowing for the investigation of the anomalous process *vis-a-vis* other components. The essence of the wavelet decomposition algorithm is that splitting of signal components is done not only low frequency domain, but also in the high frequency region. With this algorithm, the operation of splitting or decomposition is applied to any of the resulting high-frequency component, and so on down the frequency scale. Further, through the adaptive reconstruction of wavelet coefficients of the different wavelet domains containing elements of traffic anomalies, it is possible to confirm the parameters of anomalies and increase the reliability of detection. Employing wavelet packet transform method with a sliding window makes it possible to reduce computational complexity by eliminating computation redundancy The use of windows and remembering parts of the coefficients in memory effectively eliminates the need for redundant re-computations, hence speeding up the computation algorithm increasing memory usage.

The number of a_{1k} and d_{1k} coefficients is reduced by half compared to the original signal. The next iteration step for level two is executed with the approximations obtained at level 1 in a similar way. In practice, the highest level of decomposition is determined by the number n_0–1 discrete values of the signal ($N = 2^{n_0}$).As a result, at each level of *j* decomposition we have a sequence of coefficients of the approximation a_j and *detailed* d_j of length $N/2^j$ each, and the original signal can be regenerated from equation (4):

$$S(t) = a_j(t)\phi(t) + \sum_{j=1}^{m} d_j(t)\, \psi_1(t) \quad (4)$$

The number of multiplications in the direct FWT will be *2LN*, where $L = 2n$. The same number of operations is necessary for the reconstruction of the signal. Thus, for the signal analysis–synthesis in the wavelet basis, *4LN* operations must be executed, which is less than the number of operations for the fast Fourier transform ($N \log_2 N$).

A. Method

We consider the detection of network traffic anomalies based on discrete wavelet transform using statistical criteria. To adapt this method to the analysis of real-time traffic the technique of two sliding windows W_1 and W_2, moving in time with a given step is employed, while noting the value of traffic located at the time boundaries of each window.

The use of "sliding window" allows for the increase in reliability of the detection of even minor abnormalities. It is known that the spectral power density of the time series of "traffic–time", in the presence of anomalies, has peaks at a certain frequencies.Wavelet analysis allows for the detection of traffic anormalies on the basis of differences in the spectra of normal and abnormal traffic. We will consider window W_1 as '*comparison window*' and the window W_2 as a *'detection window'*. Let the size of each window W_1 and W_2 be selected time units respectively, such that $W_1 > W_2$. Then at an arbitrary time *t* the beginning of the window W_2 will be at the point *t*, and it would contain *w2* traffic values for the time interval spanning from *t–w2* to *t*. The W_1 window will contain W_1 values from *t–w2–w1* to *t–w2*.

Performing FWT for samples within each of the windows at each time t_i, we get at a certain scale level *j*, a set of coefficients $\{a_{1x}, a_{2x}, a_{3x}, \cdots, a_{nx}\}_{t,j}$ for the W_1 (approximation) window and another set $\{d_{1x}, d_{2x}, d_{3x}, \cdots, d_{nx}\}_{t,j}$ for the W_2 (detail) window; $\{a_{1y}, a_{2y}, a_{3y}, \cdots, a_{my}\}_{t,j}$ for the W_1 (approximation) window and $\{d_{1y}, d_{2y}, d_{3y}, \cdots, d_{my}\}_{t,j}$ for the W_2 (detail) window. The quality of *n* and *m* coefficients at level *j* is gotten from expressions (5) for windows W_1 and W_2 respectively:

$$n = \frac{w1}{2^j} \quad ; \quad m = \frac{w2}{2^j} \quad (5)$$

These coefficients are tested using statistical criteria, and decisions on the cardinal differences of the analyzed parameters between windows W1 and W2 will be based on the acceptance or rejection of statistical hypotheses and hence the presence of anomalies or the absence thereof will be determined. Analysis of both approximate and detailed coefficients shows that anomaly can be seen at the first level of wavelet decomposition. Therefore, FWT will be carried out on the first decomposition level, until the special statistical thresholds conditions as described below are exceeded.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

III. ANOMALY DETECTION ALGORITHM

We describe an algorithm for detecting abnormal spikes based on statistical criteria used to determine changes in the variance and the mean of the coefficients of the wavelet transform.Fisher's criterion is proposed for detecting anomalies expressed as change invariance, while the Cochran criteria is used to detect changes in the mean value [5].

The use of Fisher's criterion is proposed for detecting changes in the variances of samples of windows W_1 and W_2. The sample distribution is considered Gaussian. At any given time *t* two statistical hypothesis are proposed at scale level *j* about the equality of the variances of two samples $\{d_{1x}, d_{2x}, d_{3x}, \cdots, d_{nx}\}_{t,j}$ and $\{d_{1y}, d_{2y}, d_{3y}, \cdots, d_{my}\}_{t,j}$:

a) the null hypothesis – H_0: $\sigma^2_{1,t,j} = \sigma^2_{2,t,j}$ and

b) the alternative hypothesis – H_1: $\sigma^2_{1,t,j} \neq \sigma^2_{2,t,j}$.

The algorithm for detection of spikes in Gaussian process based on the analysis of anomalous variation of variances can be written as:

$$Z_{t,j} = \frac{S^2_{2,t,j}}{S^2_{1,t,j}} \quad (6)$$

where:

$S^2_{1,t,j} = \frac{1}{n-1}\sum_{i=1}^{n}(d_{ix} - \bar{d}_x)^2$ – sample variance of sample sequence of *details* on a scale level *j* in window W1;

$S^2_{2,t,j} = \frac{1}{m-1}\sum_{i=1}^{n}(d_{iy} - \bar{d}_y)^2$ – sample variance of sample sequence of *details* on a scale level *j* in window W1;

$\bar{d}_x = \frac{1}{n}\sum_{i=1}^{n} d_{ix}$ – sample mean of a sequence of *details* on a scale level *j* in window W1;

$\bar{d}_y = \frac{1}{m}\sum_{i=1}^{m} d_{iy}$ – sample mean of a sequence of *details* on a scale level *j* in window W2;

The use of Cochran criterion is proposed for detecting changes in the mean sample of approximations $\{a_{1x}, a_{2x}, a_{3x}, \cdots, a_{nx}\}_{t,j}$ and $\{a_{1y}, a_{2y}, a_{3y}, \cdots, a_{my}\}_{t,j}$. The algorithm for detecting spikes in traffic data based on analysis of anomalous change in sample mean values is expressed as:

$$Y_{t,j} = \frac{1}{S_{t,j}} \quad (7)$$

where:

$S^2_{1,t,j} = \frac{1}{n-1}\sum_{i=1}^{n}(a_{ix} - \bar{a}_x)^2$ – sample variance of sample sequence of *approximations* on a scale level *j* in window W1;

$S^2_{2,t,j} = \frac{1}{m-1}\sum_{i=1}^{n}(a_{iy} - \bar{a}_y)^2$ – sample variance of sample sequence of *approximations* on a scale level *j* in window W1;

$S^2_{t,j} = \frac{S^2_{1,t,j}}{n} + \frac{S^2_{2,t,j}}{m}$ – normalized sum of sample variance of *details* in windows W1 and W2;

$\bar{d}_x = \frac{1}{n}\sum_{i=1}^{n} d_{ix}$ – sample mean of a sequence of details on a scale level *j* in window W1;

$\bar{d}_y = \frac{1}{m}\sum_{i=1}^{m} d_{iy}$ – sample mean of a sequence of details on a scale level *j* in window W2;

$\bar{a}_x = \frac{1}{n}\sum_{i=1}^{n} a_{ix}$ and $\bar{a}_y = \frac{1}{m}\sum_{i=1}^{m} a_{iy}$ – sample mean of sample sequence of *details* on a scale level *j* in window W1 and W2 respectively.

Summarizing the procedure above, an algorithm for implementing the detection of anomalies based on discrete wavelet transform is hereby presented. The following actions are taken for each current window position at time t:

STEP 1. Perform Fast Wavelet Transform for 1st decomposition level on each sample from windows W1 and W2 according to equation (4);

STEP 2. Compute Fisher statistics based on the *details* coefficients d_j according to equation (6).

STEP 3. Compute Cochran statistics based on the *approximation* coefficients a_j according to equation (7).

STEP 4. Compute two thresholds for each statistic based on the accepted values of the confidence intervals with the lower threshold of p_1 = 0.95, the upper threshold p_2= 0.999.

STEP 5. Compare the current values of Fisher's and Cochran criteria with their thresholds: if either is lower than the lower threshold – go to step 6, if on the other hand, either is higher than the upper threshold – go to step 7.

STEP 6. Perform further FWT on the next decomposition level *j*. This step is only executed if the current decomposition level *j* is not greater than the maximum for the particular sequence. Repeat step 2 to step 5 for the current *j* level.

STEP 7. Reconstruct coefficients for the level at which the upper threshold was exceeded. To which end the *approximations* coefficients $A_j = a_j(t)\phi(t)$ and the *details* coefficients $D_j = d_j(t)\psi(t)$ are restored. The existence of an anomaly is documented ***only*** in the event of any of the statistical criteria exceeding the upper threshold, otherwise, there is no anomaly and the window moves on.

Thus, the ultimate test for detecting an attack is exceeding the upper threshold by one of the statistical criteria at the stage of coefficients reconstruction.

IV. DISCUSSIONS: THE DEVELOPED SOFTWARE

A software was developed in accordance with this proposed algorithm with a graphical user interface in MATLAB. The main window in the process of analyzing the sequence is shown in Figure 1. The top graph in Figure 1 shows an implementation of network traffic with attacks and the sliding

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

moving window process. The middle and bottom graphs show the Fisher and Cochran parameters calculated in real-time respectively. The red and yellow lines represent the upper and lower thresholds respectively. These graphs depict only the first decomposition level of the fast wavelet transform.

If the conditions described in step 7 of the algorithm above hold, the occurence of an attack as well as its moment of first occurrence are documented. The attacks are shown as red vertical lines in the trace (top) graph, and the number of attacks recorded in the whole sequence is displayed at the base of the GUI, in this case five attacks has been documented (shown as '5'). It can be clearly seen that the anomaly in the region of 6×10^4 is a typical DoS attack. It was well detected by both criteria (exceeds the red upper threshold) at each FWT level of decomposition. Moreover, Fisher's criterion detects this attack much more clearly, this is seen by the size of the spike and how much it exceeds the threshold of its graph.

Figure 1. Sequence Analysis Program Graphical User Interface

It is observed that majority of the anomalies occur at the initial level of decomposition 1, while some of the anomalies could have been missed if decomposition was started higher levels. We also observe that the number of false alarms are more at higher decomposition levels. This is most likely due to the low resolution of the DWT in time and consequently, small coefficients of confidence at higher levels.

A comparison of the inter-dependence of the crucial statistics shows that the determinant statistic for detecting abnormal spikes of mean value of the approximation coefficients is more efficient for Fisher's criteria than it is for Cochran. This is explained by taking into account the non-Gaussian nature of the critical statistics in the case of Fisher's criterion**.**

V. CONCLUSION

We have presented in this paper a proposed algorithm for detecting denial of service (DoS) and distributed denial of service (DDoS) attacks in information communication networks using discrete wavelet analysis. The proposed algorithm was tested by developing a software based on it in Matlab environment. Analysis of experimental results obtained using the proposed algorithm and developed software corroborates our submission on the accuracy of the proposed algorithm in detecting DoS and DDoS attacks.

ACKNOWLEDGEMENT

The authors appreciate the Lincoln laboratory of Massachusetts Institute Technology for making the (1999

DARPA Intrusion Detection Evaluation) data sets used in this study freely available on the Internet.

REFERENCES

[1] Roland Kwitt. A Statistical Anomaly Detection Approach for Detecting Network Attacks. 14th December 2004/ 6QM Workshop, Salzburg.

[2] L.Feinstein and D.Schnackenberg. Statistical Approaches to DDoS Attack Detection and Response. Proceedings of the DARPA Information Survivability Conference and Expostion (DISCEX'03), April 2003.

[3] Vinay A.Mahadik, Xiaoyong Wu and Douglas S. Reeves, "Detection of Denial of QoS Attacks Based On χ 2 Statistic And EWMA Control Charts" http://arqos.csc.ncsu.edu/papers/2002-02-usenixsec-diffservattack.pdf, NC State University, Raleigh.

[4] Nong Ye and Qiang Chen. An Anomaly Detection Technique Based on a Chi-Square Statistic for Detecting Intrusions into Information Systems. Quality and Reliability Eng. Int'l, Vol 17, No. 2, P. 105-112, 2001.

[5] E.L. Miller , "Efficient computational methods for wavelet domain signal restoration problems," Signal Processing, IEEE Transactions on , vol.47, no.4, pp.1184-1188, Apr 1999.

[6] DARPA Intrusion Detection Data Sets, Accessed: 11.01.2012, available at:http://bit.ly/xuCDby

[7] O.I. Sheluhin, A.A. Atayero, A.B. Garmashev, "Detection of Teletraffic Anomalies Using Multifractal Analysis", Proceedings of the IEEE 11th International Conference on ITS Telecommunications (ITST-2011), ISBN: 978-1-61284-670-5, DOI: 10.1109/ITST.2011.6060160, 23rd – 25th Aug. 2011, St. Petersburg, Russia.

[8] S. Mallat, "A Wavelet Tour of Signal Processing", 3rd Edition, The Sparse Way, Academic Press, USA, 2009.

AUTHORS PROFILE

Oleg I. Sheluhin was born in Moscow, Russia in 1952. He obtained an M.Sc. Degree in Radio Engineering1974 from the Moscow Institute of Transport Engineers (MITE). He later enrolled at Lomonosov State University (Moscow) and graduated in 1979 with a Second M.Sc. in Mathematics. He received a PhD at MITE in 1979 in Radio Engineering and earned a D.Sc. Degree in *Telecommunication Systems and Devices* from Kharkov Aviation Institute in 1990. The title of his PhD thesis was *'Investigation of interfering factors influence on the structure and activity of noise short-range radar'*.
He is currently Head, Department of Information Security, Moscow Technical University of Communication and Informatics, Russia. He was the Head, Radio Engineering and Radio Systems Department of Moscow State Technical University of Service (MSTUS).

Prof. Sheluhin is a member of the International Academy of Sciences of Higher Educational Institutions. He has published over 15 scientific books and textbooks for universities and has more than 250 scientific papers. He is the Chief Editor of the scientific journal *Electrical and Informational Complexes and Systems* and a member of Editorial Boards of various scientific journals. In 2004 the Russian President awarded him the honorary title '*Honored Scientific Worker of the Russian Federation'*.

Aderemi A. Atayero graduated from the Moscow Institute of Technology (MIT) with a B.Sc. Degree in Radio Engineering and M.Sc. Degree in Satellite Communication Systems in 1992 and 1994 respectively. He earned a Ph.D in Telecommunication Engineering/Signal Processing from Moscow State Technical University of Civil Aviation, Russia in 2000.
He is a member of a number of professional associations including: the Institute of Electrical and Electronic Engineers, IEEE, the International Association of Engineers, IAENG, and a professional member of the International Who's Who Historical Society (IWWHS) among others. He is a registered engineer with the Council for the Regulation of Engineering in Nigeria, COREN. He is a two-time Head, Department of electrical and Information Engineering, Covenant University, Nigeria. He was the coordinator of the School of Engineering of the same University.
Dr.Atayero is widely published in International peer-reviewed journals, proceedings, and edited books. He is on the editorial board of a number of highly reputed International journals. Atayero is a recipient of the '2009/10 Ford Foundation Teaching Innovation Award'. His current research interests are in Radio and Telecommunication Systems and Devices; Signal Processing and Converged Multi-service Networks.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

Developing an Auto-Detecting USB Flash Drives Protector using Windows Message Tracking Technique

Rawaa Putros Polos Qasha
Department of Computers Sciences
College of Computer Sciences and Mathematics
University of Mosul
Mosul, Iraq
rawa_qasha@yahoo.com

Zaid Abdulelah Mundher
Department of Computers Sciences
College of Computer Sciences and Mathematics
University of Mosul
Mosul, Iraq
zaidabdulelah@gmail.com

***Abstract* – this paper presents Windows Message Device Change Tracking (WMDCT) program to protect Windows systems from Universal Serial Bus (USB) viruses which use the AutoRun property to execute. The WMDCT program introduces a new method to develop the traditional ways of protecting techniques, which are used by other anti-viruses programs. The main two parts of WMDCT program are monitoring and tracking Windows Message Device Change, which is a message that is sent by the system, in the background, and removing or repairing the infected files in the USB flash drive. WMDCT has been tested in the University of Mosul/ Computer Science Dept. labs and the results have been mentioned in this paper.**

Keywords-USB; AutoRun; system protection;Windows Messages

I. INTRODUCTION

Universal Serial Bus (USB) storage devices are one of the most common means of viruses to attack computers. Nowadays, there are many viruses exploit the lack of security mechanism for Windows Autoplay features to attack Windows systems. According to McAfee Avert Labs [1], the top rank of Malware is AutoRun Malware. In addition, according to Ghosh [2], half of the top 10 viruses of 2009 exploited the Windows AutoRun feature. The WMDCT introduces a new, fast, and efficient approach to protect Windows systems from viruses' infection which are used USB flash drive with AutoRun property to separate. The WMDCT approach depends on tracking the WM_DEVICECHANGE message, which is sent by the Windows system to all applications when a USB device connects to the system. When WMDCT program receive this message, it checks if the flash drive contain an AutoRun.inf file to be removed, which makes the viruses files completely paralyzed. WMDCT program also restores the default properties of the other files that have been infected by the virus. This method has been provided the following features:

- Removing the AutoRun.inf file automatically in a non interactive way makes the WMDCT program very useful with computers which are used by different users such as in computers labs at universities.
- Removing a specific file (AutoRun.inf) makes the update process not necessary.
- Removing only the AutoRun.inf file, which is put on the root of the flash drive, makes the WMDCT program very fast.

II. RELATED WORKS

Some related work such as Wolle, J., suggested stopping AutoRun property from the Control Panel [3]. Clearly, this is not a real solution because if the user pressed double-click to open the USB flash drive, the system will be infected since the AutoRun.inf file still on the USB flash drive. To the best of the researcher's knowledge, this solution to protect computers from AutoRun malware attacks has never been used or posed before. According to Aycock, J., the first task of anti-virus programs is detecting if other programs are a virus or not [4]. There are many algorithms which are used for this purpose such as Aho-Corasick, Veldman, and Wu-Manber. These algorithms depend on set of signatures to detect viruses. Traditionally, anti-virus programs use signatures to identify viruses. The two major disadvantage of this method are that it needs new signatures to detect new viruses, and it is slow down the system since it uses complex algorithms. All the related works try to enhance those methods to reduce amount scans and resource requirements. The Pham, D., Halgamuge, M., Syed, A., Mendis, P. introduced a new method also using AutoRun file to protect only USB flash drives not the computers [5]. The aim of this work is to introduce a simple but efficient method to protect Windows systems from AutoRun viruses/malwares.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

III. AUTORUN FILE AND WM_DEVICECHANGE MESSAGE

A. According to Szor, P., AutoPlay is the feature built into Windows that automatically runs a program specified by the file AutoRun.inf whenever a CD-ROM, DVD or USB drive is plugged into a Windows-based computer [6]. Moreover, Tahir, R., Hamid, Z., Tahir, H., noted that "Flash drive infections usually involve malware that loads an AutoRun.inf file into the root folder of all drives (internal, external, and removable) which automatically runs a malicious .exe file on the computer [7]. When an infected USB flash drive is inserted, the Trojan infects the system." The Autorun section supports an open command that can be used to run executable files. This is the command that malicious codes exploit to be invoked automatically. A simple Autorun.inf file is:

[autorun]
open=autorun.exe
icon=autorun.ico

B. According to Microsoft Developer Network [7] and Axelson, J. [8], Windows sends all top-level windows a set of default WM_DEVICECHANGE messages when new devices or media (such as a CD or Flash Drive) are added and become available. When the user inserts a new CD, DVD, or Flash drive, applications receive a WM_DEVICECHANGE message with a DBT_DEVICEARRIVAL event. DBT_DEVICEARRIVAL is sent after a device or piece of media has been inserted. Applications receive this message when the device is ready for use as kind of notification. Each notification contains a device path name that the application can use to identify the device that the notification applies to.

IV. PROPOSED METHODOLOGY

The main advantage of this work is that the removed operation will be applied in the background without user interaction. When a USB flash drive connects to the computer, WMDCT will discover it automatically and remove the malicious files from it. As mention previously, when a USB device connects to a computer, the Windows system sends the WM_DEVICECHANGE message to applications. WMDCT starts with listening to this message. As soon as WMDCT receives WM_DEVICECHANGE message, the scan operation on the connected device is performed. If WMDCT detect any AutoRun.inf file in the connected USB flash drive, WMDCT will change the permission of it to normal and removed it. Also, depending on settings that the user are selected from the WMDCT interface, all the EXE files or the EXE files with hidden attribute will be removed. Another feature which WMDCT introduced is that using multi-threading technique to improve the performance of the WMDCT. Sometimes more than one USB flash drive connects to the computer at the same time which causes an overlap. This problem has been solved by using multi-threading technique by create a separated thread for each new USB flash drive which connects to the computer. The following flowchart demonstrates the algorithm which is implemented by WMDCT program to protect Windows systems from viruses that execute using AutoRun property.

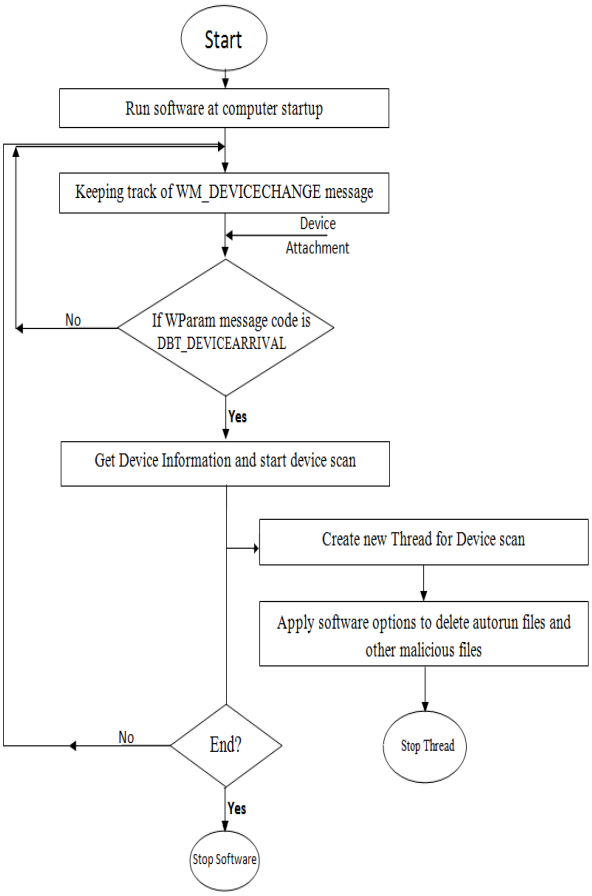


Figure 1: WMDCT algorithm

V. EXPERIMENTS AND DISCUSSION

C# language with .NET 4.0 platform was used to develop WMDCT program. WMDCT program was tested in the University of Mosul/ Computer Science

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

Dept. Labs and many other personal computers. The results have shown the efficiency of WMDCT. The most important features which are provided by WMDCT are speed and independence. WMCDT was tested on computers which are used by many different users (students), and each student has different USB flash drive. WMDCT was very efficient and the percentage of success to delete AutoRun.inf files was 100%. Figure (1) shows WMDCT interface which gives the administrator/user the ability to set up the program options.

Figure2: WMDCT main interface

Table (1) explains these options**.**

Table (1): WMDCT options

Option	Function
Remove autorun.inf file	Remove the AutoRun.inf file automatically.
Remove all EXE files in root on Removable disk	Removes all execution files in the root directory of the detected USB flash drive.
Remove only XE file with hidden attribute	Removes only hidden execution files in the root directory of the detected USB flash drive.
Show hidden files and directories on Removable disk	Show all the hidden files and directories which are mostly expected to be infected by viruses.
Run program with startup	Run WMDCT automatically when Windows startup.

VI. EVALUATION AND COMPARISON

The system was evaluated by monitoring the time and the CPU usage. Figure (3) and Figure (4) show the results of this evaluation:

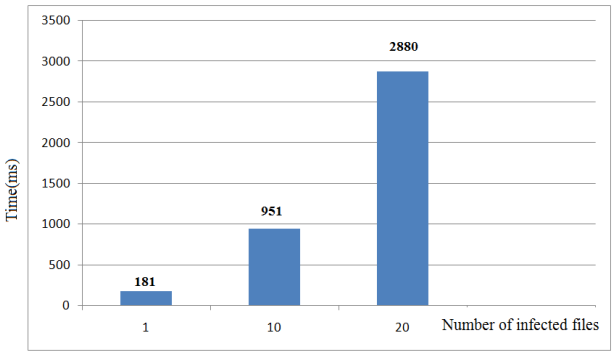


Figure 3: Time measurement

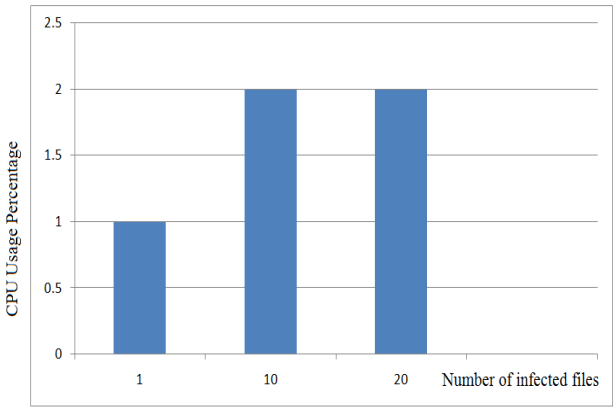


Figure 4: CPU usage measurement

In addition, Table (2) shows a comparison between traditional anti-virus programs and WMDCT program.

Table 2: the comparison between anti-virus programs and WMDCT program

	Other anti-virus programs	WMDCT
System Performance	Adversely affect in different proportions	No significant effect
Speed	Scanning need a long time	Very fast
Update	Require an up-to-date database of virus signatures	No update is required
Efficiency	Only Known viruses are detected	Known and unknown viruses are

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

		detected
Detection	All types of viruses are detected	Only AutoRun viruses are detected

Moreover**,** According to Aycock, J. [4], there are some sophisticated viruses use anti-anti-virus techniques to avoid detection by anti-virus programs. Up until now, there is no one of these techniques can pass the WMDCT program since viruses use these techniques trying to make analysis difficult for anti-virus programs, while WMDCT do not try to analyze viruses' files. WMDCT try to stop the mechanism which is used by viruses to execute, which is represented by AutoRun.inf file.

VII. CONCLUSIONS

There are many serious threats associated with the use of USB flash drives, and many of these threats depend on AutoRun mechanism to execute. This paper suggested and implemented a new solution to protect computers from this kind of viruses by introducing WMDCT program to detect any connection with USB flash drives and remove the AutoRun.inf file automatically. This solution does not require complex configuration or high system resources. Windows messages are the magic key that was used to achieve this work.

REFERENCES

[1] McAfee Avert Labs., *"McAfee threats report: Second quarter"*, McAfee, Inc., 2011.

[2] Ghosh, A. *"Ten Most Threatening Viruses of 2009"*. Retrieved Nov. 26, from http://www.brighthub.com/computing/smb-security/articles/44811.aspx, 2011

[3] Wolle, J., *Malware Protection White Paper, 2006.*

[4] Aycock, J., "Computer Viruses and Malware. Canada". Springer, 2006.

[5] Pham, D., Halgamuge, M. , Syed, A., Mendis, P, "*Optimizing Windows Security Features to Block Malware and Hack Tools on USB Storage Devices*", *PIERS Proceedings*, 350-355, 2010.

[6] Szor, P., "*The Art of Computer Virus Research and Defense"*, Addison Wesley Professional, 2005.

[7] Tahir, R., Hamid, Z. , Tahir, H., "*Analysis of AutoPlay Feature via the USB Flash Drives*", *World Congress on Engineering*, Vol I., 2008.

[8] Axelson, J. *"USB Complete: The Developer's Guide"*, 4th Edition, 2009.

AUTHOR PROFILE

Miss Rawaa P. Qasha (MSc.) is currently a lecturer at Mosul University/ College of Computer Science and Mathematics/ Computer Science Department. She received B.Sc. degree in Computer Science from University of Mosul in 1997 and M.Sc. degree from University of Mosul in 2000. Her research interests and activity are in operating system, operating system security, distributed systems, mobile operating system, virtualization, and computer clouding. Now, she teaches Operating System and Programming Languages for undergraduate students.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

Analysis of DelAck based TCP-NewReno with varying window size over Mobile Ad Hoc Networks

Parul Puri[1] Gaurav Kumar[2] Bhavna Tripathi[3]

Department of Electronics & Communication Engineering
Jaypee Institute of Information Technology,
Noida, India.
parulpuri9@gmail.com[1]
er.gauravchachra@gmail.com[2]
my.bhavna@gmail.com[3]

Dr Gurjit Kaur[4]

Assistant Professor,
Department of Electronics & Communication Engineering
School of ICT,
Gautam Buddha University,
Greator Noida, India.
gurjeet_kaur@rediffmail.com[4]

***Abstract*—In this paper, we study TCP performance over multi-hop wireless networks that use IEEE 802.11 protocol for access. For such networks NewReno is the most deployed TCP variant that handles multiple packet losses efficiently. It is shown that the delayed ACK scheme substantially increases the TCP throughput. We propose an approach to improve the performance of half-duplex and asymmetric multi hop networks widely employed for mobile communication. Our approach is based on optimizing the timer duration of the delayed ACK scheme and varying the window size. Simulations have been carried on NS2 for TCP-NewReno variant using DSDV and AODV routing protocols.**

Keywords: Multi-hop wireless networks, TCP, Newreno, DelAck, DSDV, AODV.

I. INTRODUCTION

In the last few years, many research works have focused on multi-hop wireless networks, in which relaying nodes are in general mobile, and communication needs are primarily between nodes within the same network. In such networks, a number of intermediate nodes whose function is to relay information from one point to another point carry out communication between the two end nodes. The application can be useful in various fields, especially because it uses wireless means of communication, hence saving the hassle of laying down wires in already crowded or remote terrains. People working in collaboration and places in remote locations can connect through it. Activities which require working at locations having no ground infrastructure, like patrolling, disaster hit areas and rural areas, can be carried out using this technology. Some important applications are also being developed on the basis of this technology which can be used by armed forces in rescue and war time scenarios [1].

Two key requirements of any network are reliable data transfer and congestion control. The transmission control protocol (TCP) was designed to provide reliable end-to-end delivery of data packet in the wired networks. However, unlike wired networks wireless networks suffer from many problems, such as packet losses due to congestion, node mobility, high bit errors, medium access contention due to hidden terminals, and so on. Hence, in order to apply TCP in a wireless environment, TCP needs some modifications. Further, keeping in mind the basic characteristic of a TCP scheme the acknowledgement (ACK) packets need to be transmitted from TCP sink to TCP source, against the flow of TCP data packets. This results in simultaneous arrival of TCP data and ACK packets which can cause collisions and even packet losses [2, 3]. As a result, there is a huge degradation in throughput in multi-hop networks [4].

At the MAC level, each data packet transmission is a part of four-way handshake protocol, which is intended to reduce the collision probability. The handshake reduces the probability of hidden-terminal collisions, but it does not eliminate them. This limits the number of packets that can be transmitted simultaneously in a wireless network without collisions. The main factor affecting the TCP performance in multi-hop wireless networks is the contention and collision between ACK and data packets caused by taking the same path. Thus, in order to improve the TCP throughput, we shall try to decrease the ACK flows by using the delayed ACK scheme, where an ACK is transmitted for every d packets, defined by the DelAck number, that reach the destination [5]. However, to avoid a deadlock, and if d packets do not arrive, an acknowledgement is generated after some time interval without further waiting.

The throughput of a network is limited by two windows: the congestion window and the receive window. The TCP sender uses a congestion window (cwnd) in regulating its

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

transmission rate based on the feedback it gets from the network [6]. Whereas, the receive window size sets a limit on the amount of data that can be sent unacknowledged. Earlier researches on TCP performance over multi-hop wireless networks [3] have shown that for static chain topology it is beneficial to limit the maximum receive window size of TCP sink to around n/4, where n is the number of nodes; and any further increase in the maximum window size causes more collisions and deterioration in the throughput. However, the issue of limit on an optimum window size for mobile topology is left unaddressed.

It is also seen, for a fixed small size of maximum window size, the delayed ACK does not outperform the standard TCP version since most of the time, the window size limits the number of packets that can be transmitted by the sender to less than d. So, the delayed ACK scheme has to wait for the timer to expire before generating an ACK; and the sender cannot transmit packets during that time. Hence, the time interval plays a critical part of TCP system with DelAck scheme.

Tahiliani et al in [4] has studied the performance of TCP variants such as Tahoe, Reno, NewReno, Sack, and Vegas over various routing protocol. They have analyzed that TCP NewReno and Sack perform better in comparison to the other schemes. In this paper, the NewReno variant of TCP is tested as it is the most deployed one. We propose an approach to improve the TCP performance by simulating the delayed ACK scheme with an optimum time interval and by varying the receive window size for the same size of congestion window (cwnd) for mobile topology. We choose one proactive routing protocols: Destination Sequenced Distance Vector (DSDV) as well as one reactive routing protocols: Ad hoc On demand Distance Vector (AODV) for our study since they are accepted as the standard routing protocols for multi-hop wireless networks [7].

II. Related Work

In Reference [2], G. Holland et al uses a new metric called *expected throughput* to compare the performance by measuring the differences in throughput with varying number of hops. Further the authors have studied the effects of mobility on TCP Reno's performance in mobile ad hoc networks. This metric will be used in our paper and will be discussed in detail in Section V.

Ammar Mohammed AI-Jubari [5] has shown that the delayed acknowledgment strategy can improve TCP throughput up to 233% compared to the regular TCP over multi-hop wireless networks.

Jiwei Chen [8] has tried to explain the effect of receive window size on the TCP throughput, but have restricted the research to static topology only.

III. Delayed ACK Scheme

RFC 831 first suggested a delayed acknowledgement (DelACK) strategy, where a receiver doesn't always immediately acknowledge segments as it receives them. This recommendation was carried forth and specified in more detail in RFC 1122 and RFC 5681 (formerly known as RFC 2581). RFC 5681 mandates that an acknowledgement be sent for at least every other full-size segment, and that no more than 500ms expire before any segment is acknowledged.

Basically, the delayed acknowledgement procedure defines two terms: DelAck number and Time interval. The DelAck number d defines the number of packets for which the receiver waits before sending an acknowledgement. By using delayed acknowledgement mechanism the numbers of acknowledgments required are reduced. As acknowledgments are also parts of traffic, the load over channel decreases. Thus, using this concept the throughput is increased. But this is not always the case; there are some situations where delayed acknowledgment leads to reduction in bandwidth. Studies have shown d = 2 gives an optimum performance.

Second parameter of the delayed acknowledgement procedure is the Time Interval (Fig. 1). A timer is set by the TCP, depending on which DelAck procedure is modified. Now the acknowledgement is sent when the two packets are received or if the timer goes off, whichever occurs first.

We aim to study the effect of the delayed acknowledgement procedure on TCP throughput over multihop wireless links.

Jiwei Chen et al [8] has studied that increasing the value of DelAck number does not always show a positive increase in the throughput. In some situations it has proved to be deteriorating also. This is so because if a large DelAck number is chosen it will cause a large burst of packets to pass thereby increasing interference. Keeping in view this adverse affect we have kept our DelAck number to be 2 and focus our study on the Time interval aspect.

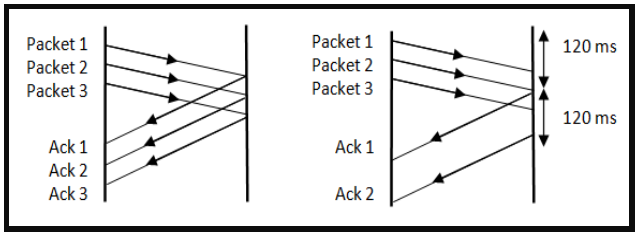


Figure 1. Role of DelAck and Time Interval in TCP communicattion

IV. Window Size

In order to limit the impact of congestion, TCP uses a special kind of buffer called Sliding (Receive) Window. Receive window size indicates the buffer size of the receiver. In other words, window size is the maximum number of packets (bytes) a source can transmit before receiving an acknowledgement from the receiver. By controlling the window size, a receiver can control the rate at which other hosts send data to it. For the small window size, the number of packets transmitted to the receiver is less. But the number of acknowledgements transmitted in this case will be comparatively larger and will cause collision with data packets, thus reducing the throughput. On the other hand, if the window size is too large, number of acknowledgements decrease. However, as the receiver buffer size is more, number of packets transmitted by the sender host increases thereby causing bursty traffic. This causes interference and packet losses depending upon the path length. Thus, there exists an

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

optimum window size for which the channel gives maximum throughput. We aim to find the size of this optimum size of the Sliding window.

V. Simulation Setup and Methodology

Simulations have been done on ns-2 [9], a discrete event simulator. The simulations were carried for multihop wireless static and mobile topologies.

A. Multihop Wireless Static Topologies

A linear string topology of 8 nodes was designed, similar to the one used in [10]. A single TCP connection with variable number of hops (1-7) was studied. The nodes were configured to use 802.11 MAC protocol with the following parameters. Distance between two nodes was 250 metres. This distance is same as the maximum transmission range. Radio propagation model used was Two-ray ground reflection model. The channel data rate was 2 Mbps, TCP packet size was 1460 bytes and the maximum window size was 32. With the above mentioned parameters fixed and varying the TCP protocol, routing protocol and TCP sink results were taken. The results have been discussed in Section VI.

B. Multihop Wireless Mobile Topologies

Our network model constitutes of 25 nodes in a 1500 x 400 m2 flat, rectangular area. Movement of nodes was according to the mobility patterns generated by the mobility pattern generator offered by ns-2; which is based on random waypoint mobility model. In this model, each node picks a random destination. Once it arrives to the destination it pauses for some time and then picks another destination. This procedure is followed throughout. The mean speed of the nodes was taken 10m/s and the pause time was 0 sec. The simulation results are based on an average throughput of 25 mobility patterns. The parameters were same as those taken for static topologies. Here, the TCP-NewReno variant was studied with variations in TCP sink, routing protocol and window size. Simulation results are discussed in the Section VI.

C. Performance Metric

Throughput has been used as the performance metric. Throughput was measured for fixed sender and receiver nodes over the entire period of the connection. TCP cannot determine the cause of packet loss, and considers congestion the reason behind the losses. Thus, the throughput so obtained is always less than the optimal value. In order to compare the difference, we use another metric called the *expected throughput*. *Expected throughput* gives an upper bound on the TCP throughput. *Expected throughput* is calculated using the throughput values obtained in the static topologies. If t_i = time, T_i = throughput, where i = hops (ranges from 1 to 7). Hence t1 means "amount of time source and destination were 1 hop far from each other". Similar explanation comes for throughput. T_2 means "throughput when source and destination were 2 hops far from each other". The values of T_i are those obtained from simulating static topologies and t_i is obtained from the scene file. So, we calculate the *expected throughput* using (1) as follows:

$$\text{Expected throughput} = \frac{\sum_{i=1}^{\infty} t_i \times T_i}{\sum_{i=1}^{\infty} t_i} \quad (1)$$

Practical Throughput is obtained from the simulations. Both *expected* and practical throughputs are then compared in terms of the percentage achieved of the *expected throughput* calculated as follows:

$$\text{Percentage Achieved} = \frac{\text{Practical Throughput}}{\text{Expected Throughput}}\% \quad (2)$$

VI. Results and Analysis

A. Multihop Wireless Static Topologies

Tables I and II show the throughput (in Kbps) obtained for each variant of TCP with DSDV and AODV routing protocols respectively. These results will be used for calculating the *expected throughput* values as explained in Section V.

Our studies show that NewReno variant of TCP gives the most optimum performance as compared to other variants for both the routing protocols. This is because of the fact that NewReno is more capable in handling multiple packet losses from a single window of data as compared to other TCP variants. Hence, for mobile topologies we carry out our analysis for the NewReno TCP scheme.

As is known, the performance of TCP depends on the routing protocols as every routing protocol has a different technique to handle link failures and to form routes. From our results, it can be seen in static topologies performance of proactive routing protocol (DSDV) is better in terms of throughput as compared to reactive routing protocol (AODV). The reason is that proactive protocols maintain a routing table. However, in reactive protocols route calculation is on-demand basis which causes some delay in sending data. Also, DSDV has lesser number of control packets which decreases the number of collisions.

Further, an improvement in throughput is observed when DelAck is used for all TCP variants over DSDV and AODV routing protocols.

B. Multihop Wireless Mobile Topologies

Tables III and IV show the throughput (in Kbps) obtained for the NewReno variant of TCP with DSDV and AODV routing protocols respectively. Throughput values have been obtained by varying the characteristics of TCP sink such as window size and delay interval.

Based on the simulation results Fig. 2 to Fig. 7 have been plotted and will be further analyzed.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

TABLE I. THROUGHPUT (IN KBPS) USING DSDV

No of Hops	Tahoe		Reno		New Reno		Sack	
	Without DelAck	*With DelAck*	*Without DelAck*	*With DelAck*	*Without DelAck*	*With DelAck*	*Without DelAck*	*With DelAck*
1	752.19	802.40	752.19	802.40	752.19	802.39	752.19	802.39
2	376.60	402.15	376.60	402.15	376.60	402.15	376.60	402.15
3	251.15	271.74	224.98	271.74	224.98	271.74	165.08	271.74
4	173.44	185.36	164.70	180.06	160.00	185.58	179.79	184.50
5	152.62	164.44	140.10	159.88	155.98	121.59	154.48	160.52
6	141.22	148.07	124.32	143.43	143.05	152.84	144.65	151.25
7	133.16	139.06	123.75	131.73	135.36	148.58	74.26	79.77

TABLE II. THROUGHPUT (IN KBPS) USING AODV

No of Hops	Tahoe		Reno		New Reno		Sack	
	Without DelAck	*With DelAck*	*Without DelAck*	*With DelAck*	*Without DelAck*	*With DelAck*	*Without DelAck*	*With DelAck*
1	757.76	805.10	757.76	805.10	757.76	805.10	757.76	805.10
2	379.15	403.50	379.15	403.50	379.15	403.50	379.15	403.50
3	198.02	222.21	199.60	222.21	211.56	222.21	203.98	217.61
4	151.24	178.50	127.64	154.55	152.46	177.88	150.65	174.58
5	127.37	152.30	113.98	137.77	130.05	152.47	126.17	150.17
6	116.77	136.02	105.44	125.04	119.80	135.58	118.47	133.21
7	51.81	75.06	53.89	99.08	56.25	71.39	42.29	107.01

TABLE III. THROUGHPUT (IN KBPS) USING DSDV

Window Size	Without DelAck	DelAck-100 ms	DelAck-120 ms	DelAck-140 ms
2	496.36	520.60	535.20	523.08
4	509.72	531.44	535.68	532.44
6	491.44	524.64	528.64	538.68
8	525.77	564.11	546.54	559.28
20	519.84	560.02	544.97	547.22
32	524.11	560.76	570.00	549.33
Expected Thpt	592.48	634.31	634.31	634.31

TABLE IV. THROUGHPUT (IN KBPS) USING AODV

Window Size	Without DelAck	DelAck-100 ms	DelAck-120 ms	DelAck-140 ms
2	513.96	539.28	525.88	538.76
4	498.84	549.08	544.08	534.04
6	504.60	555.96	540.68	546.08
8	501.68	554.44	543.28	542.80
20	514.80	559.24	537.44	543.28
32	503.16	554.44	547.52	548.12
Expected Thpt	595.17	632.91	632.91	632.91

From Fig.2 it is seen that DSDV gives a maximum throughput of 570 Kbps for window size of 32 and a delay of 120 ms. In this case 90% of the expected throughput is achieved. However, for other delay intervals (0, 100, and 140 ms) window size-8 outperforms all other window sizes including 2, 4, 6, 20, and 32. The percentage achieved is 89% of expected throughput for window size - 8.

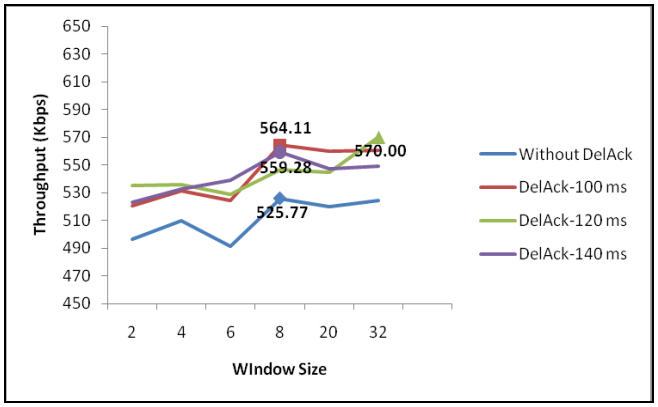


Figure 2.Throughput (in Kbps) using DSDV with varying Window Size

For window sizes 2, 4, and 6 the throughput is lesser than the optimum window size - 8 for different delay intervals. This decrease in throughput for small window sizes at higher intervals is evident as for small window sizes the buffer capacity of receiver is small. Hence, the sender can now send a limited number of packets until it has received acknowledgements for all packets in that window. However, as the timer interval is more, receiver remains idle for a longer duration before sending the acknowledgement. This results in a decrease in throughput. On similar grounds, the adverse effects of elevated idle time are observed for 140 ms delay interval for all window sizes. This indicates the limitation on the value beyond which delay interval should not be increased.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

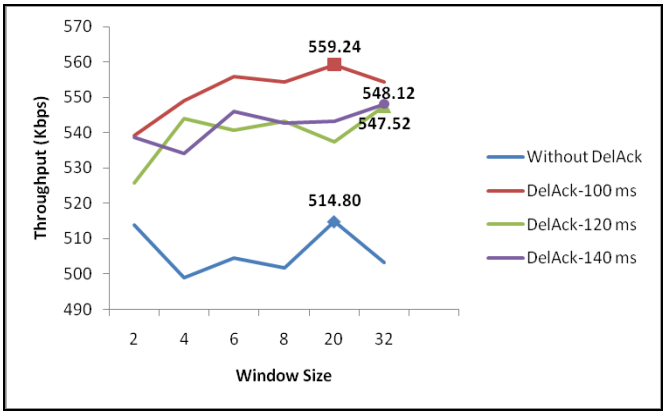


Figure 3. Throughput (in Kbps) using AODV with varying Window Size

In case of AODV, as seen from Fig. 3, peak in throughput obtained is 559 Kbps at a window size of 20 with delay 100 ms. It has achieved 88% of the expected throughput. In comparison to the DSDV protocol, AODV has some variations in terms of the optimum window size and delay interval. As seen from Fig. 3, peaks in throughput values are obtained for larger window sizes such as 20 and 32 for different time intervals, in comparison to DSDV where the optimum size of window for different time intervals was 8. In terms of the delay interval, Fig. 5 shows that best performance for AODV is obtained for DelAck=100ms. Any further increase in the delay interval degrades its performance. Overall, performance of DSDV is better in comparison to the AODV protocol.

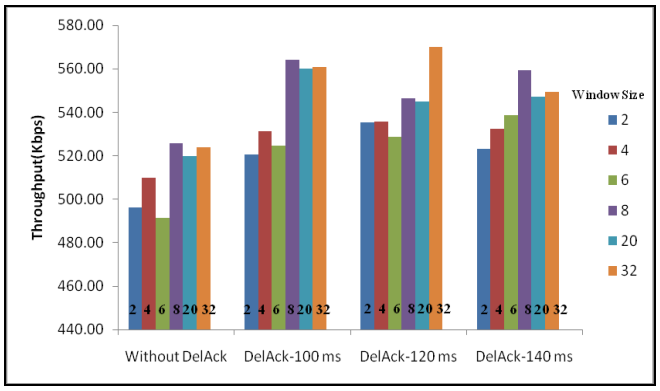


Figure 4. Throughput (in Kbps) using DSDV with varying Window Size and Time Intervals

Further, from Fig. 4 we analyze the gain in the throughput values obtained with DelAck and without DelAck. As expected theoretically, a significant amount of gain is obtained using DelAck. Also, the amount of gain is dependent on the two parameters, delay and window size. For smaller window sizes (2 and 4) the gain is more for 120 ms delay. For eg. for window size 4, throughput gain is 39 Kbps for 120 ms delay while for delays 100 and 140 ms the gain is 24 and 27 Kbps respectively. For larger window sizes (8 and 20) a delay of 100 ms gives the highest throughput. In case of AODV, Fig. 5 shows maximum gain is achieved for 100 ms delays for all window sizes. For 100 ms delay, gains as high as (25, 50, 51, 53, 44, 51) Kbps are obtained for window sizes 2, 4, 6, 8, 20, and 32 respectively in comparison to gains (12, 45, 36, 42, 23, and 44) Kbps for 120 ms delay.

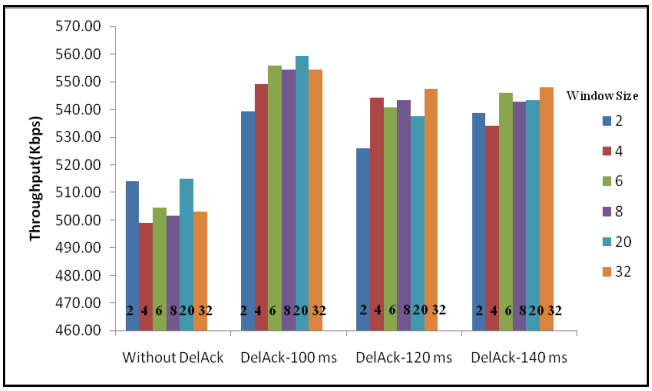


Figure 5. Throughput (in Kbps) using AODV with varying Window Size and Time Intervals

Fig. 6 gives a comparison of the expected throughput values and the practical throughput values obtained through simulations. The practical throughput values taken for comparison are the maximum values obtained for respective time intervals (100, 120, and 140 ms). It is seen, in order to achieve practical throughput values as close to the expected throughput, it is important to select time interval in conjunction with the window size.

Fig. 7 gives the values of the respective time intervals for different window sizes (2, 4, 6, 8, 20, and 32) which give maximum throughput. As can be seen for DSDV, window sizes 8 and 20 give maximum throughput of 564 Kbps and 560 Kbps at 100 ms time interval. For AODV all window sizes give maximum throughput at 100 ms time interval.

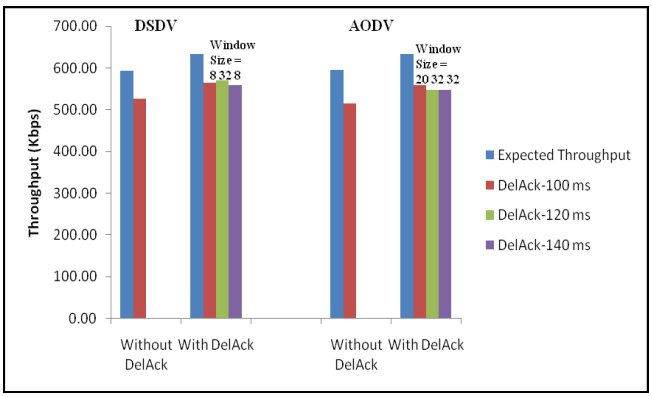


Figure 6. Comparison of Expected and Maximum Practical Throughput (in Kbps) using DSDV and AODV with varying Window Size and Time Intervals

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

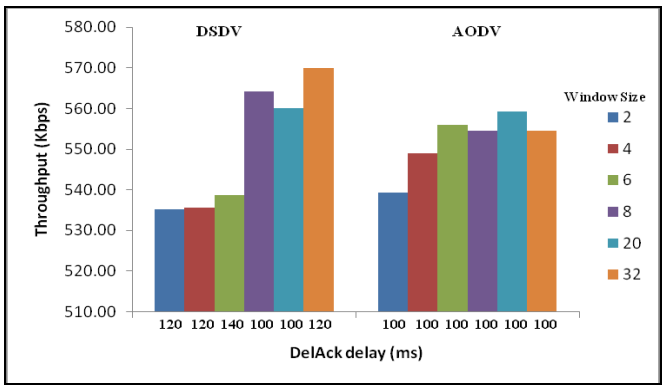


Figure 7. Plot of maximum throughputs (in Kbps) obtained for different window sizes with their delay intervals using DSDV and AODV

VII. Conclusions and Future Work

Through simulation we have studied the effect of delayed acknowledgment with variations in time interval for various receive window sizes on TCP NewReno in mobile multi-hop wireless networks. It is evident from the results that, there exists a tradeoff between the time interval and window size. We propose that maximum throughput can be achieved by selecting an optimum time interval for a particular window size. Further, it is seen choice of window size and time interval varies with the routing protocols also. Results show for DSDV a time interval of 120 ms and a large window size of 32 gives a peak in throughput. However, a window size of 8 gives the most optimum results for various delay intervals. In case of AODV, 100 ms delay with variable window size gives optimum throughput.

Currently, we are also analyzing the effect of number of nodes on the choice of window size and time interval. Testing our approach in a real test-bed experiment, to show its efficiency in the real TCP, is a part of our future work.

References

[1] M. Gerla and J.T.-C. Tsai, "Multicluster, mobile, multimedia radio network," ACM/Baltzer Journal of Wireless Networks, vol. 1, no. 3, pp. 255-265, 1995.

[2] G. Holland and N. Vaidya, "Analysis of TCP performance over mobile ad hoc networks," in Proceedings of ACM/IEEE MOBICOM, Seattle, Washington, August 1999 .

[3] T. Kuang, F. Xiao, and C. Williamson, "Diagnosing wireless TCP performance problems: a case study," in Proceedings of SCS SPECTS Conference, Montreal, PQ, pp. 176-185, July 2003.

[4] M. Tahiliani, K.C. Shet, and T.G. Basavaraju, "Performance evaluation of TCP variants over routing protocols in multi-hop wireless networks," ICCCT'10.

[5] A.M. Al-Jubari and M. Othman, "A new delayed ACK strategy for TCP in multi-hop wireless networks," Information Technology (ITSim), pp. 946 – 951, June 2010 .

[6] Pasi Sarolahti, "Linux TCP," Nokia Research Centre.

[7] Z. Fu, P. Zerfos, H. Luo, S. Lu, L. Zhang, and M. Gerla, "The impact of multi-hop wireless channel on TCP yhroughput and loss," in Proceedings of IEEE INFOCOM, San Francisco, CA, April 2003.

[8] E. Jiwei Chen, Yeng Zhong Lee, Mario Gerla, and M.Y. Sanadidi, "TCP with delayed ack for wireless networks," in Broadband Communications, Networks and Systems, pp. 1-10, October 2006.

[9] K. Fall, K. Vardhan, "The ns manual," The VINT Project, January 2009.

[10] M. Gerla, K. Tang, R. Bagrodia, "TCP performance in wireless multihop networks," in Proceedings of IEEE WMCSA, New Orleans, LA, February 1999.

AUTHORS PROFILE

Parul Puri received the B.Tech degree in Electronics & Communication Engineering from National Institute of Technology, Hamirpur, H.P., India. She is currently pursuing the M.Tech degree in Electronics & Communication Engineering from Jaypee Institute of Information Technology, Noida, India.

She has worked as a Patent Analyst in a leading legal process outsourcing firm CPA Global, Noida. She has hands on experience in patent analysis, patent infringement, and patent portfolio management in various technology domains including 'Speech', 'IP Multimedia Subsystem architecture', and 'Biometrics'.

Her current research interests include spread-spectrum communication, multi-carrier communication, channel coding, and channel fading.

Gaurav Kumar received the B.Tech degree in Electronics & Communication Engineering from Kurukshetra University, Kurukshetra, India in 2008 and is currently pursuing the M.Tech degree in Electronics & Communication Engineering from Jaypee Institute of Information Technology, Noida, India.

His current research interests include digital and wireless communication, resource allocation for broadband wireless transmissions, simulation of telecommunication systems and image processing in VHDL.

Bhavna Tripathi received the B.Tech degree in Electronics & Communication Engineering from Gautam Buddh Technical University, Lucknow , India in 2010 and is currently pursuing the M.Tech degree in Electronics & Communication Engineering from Jaypee Institute of Information Technology, Noida, India.

Her current research interests include digital and wireless communication, digital signal processing, simulation of telecommunication systems and radio-navigation systems.

Dr Gurjit Kaur has been an Assistant Professor with the Gautam Buddha University, Greater Noida, India. She received her ME and Ph.D degrees both from the PEC University of Technology, Chandigarh in 2003 and 2010 respectively. She has been a topper throughout her academic career and has received the gold medal from Honorable President of India for being overall topper at Punjab Technical University, Jalandhar.

Her professional research areas are Wireless and Optical Communication. She has many research papers of national and international repute to her credit. She has served as a reviewer of journals and conferences.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

Distributed Intrusion Detection System for Ad hoc Mobile Networks

Muhammad Nawaz Khan[a]
School of Electrical Engineering & Computer Science,
National University of Science & Technology (NUST)
Islamabad, Pakistan.
[a] (09msccsnkhan@seecs.edu.pk/nawazpk805@gmail.com)

Muhammad Ilyas Khatak[b]
Department of Computing,
Shaheed Zulfikar Ali Bhutto Institute
Of Science & Technology Islamabad, Pakistan
[b] (uomian_888@yahoo.com)

Ishtiaq Wahid[c]
Department of Computing & Technology,
Iqra University Islamabad
Islamabad, Pakistan
[c] (ishtiaqwahid@iqraisb.edu.pk)

Abstract- *In mobile ad hoc network resource restrictions on bandwidth, processing capabilities, battery life and memory of mobile devices lead tradeoff between security and resources consumption. Due to some unique properties of MANETs, proactive security mechanism like authentication, confidentiality, access control and non-repudiation are hard to put into practice. While some additional security requirements are also needed, like co-operation fairness, location confidentiality, data freshness and absence of traffic diversion. Traditional security mechanism i.e. authentication and encryption, provide a security beach to MANETs. But some reactive security mechanism is required who analyze the routing packets and also check the overall network behavior of MANETs. Here we propose a local-distributed intrusion detection system for ad hoc mobile networks. In the proposed distributed-ID, each mobile node works as a smart agent. Data collect by node locally and it analyze that data for malicious activity. If any abnormal activity discover, it informs the surrounding nodes as well as the base station. It works like a Client-Server model, each node works in collaboration with server, updating its database each time by server using Markov process. The proposed local distributed- IDS shows a balance between false positive and false negative rate. Re-active security mechanism is very useful in finding abnormal activities although proactive security mechanism present there. Distributed local-IDS useful for deep level inspection and is suited with the varying nature of the MANETs.*

KEYWORD: MANETs, Intrusion Detection System (IDS), security mechanism, proactive, reactive, Markov process, false negative and false positive.

I. INTRODUCTION

MANETs is an autonomous system of mobile nodes, built on ad hoc demands and work as wireless network, nodes move from place to place in peer to peer fashion. MANET has no pre-define structure, no centralized administration, hence any node may leave or enter the network. The self organizing nature of the ad hoc network comprises the nodes into arbitrary and temporary ad hoc topology, this leads to inherent weakness of security [1]. Security for an infrastructure-less and ad hoc nature of the network is a great challenged. On the other hand the resources constraints (limited power, limited communication range, processing capabilities, and limited memory) of the mobile devices in the MANET leads trade off s between security requirements and resources consumptions [2].

Most of the time security in ad hoc network ensures by using encryption and authentication. But the changing topology and decentralized management of MANETs, mobile nodes are compromised in many ways. Actually these protocols do not examine the received packets and do not analyze the overall network behavior but works in a traditional proactive manner. Therefore another reactive mechanism is required which not only check the packets locally but also deeply inspect that what is the internal state of the receiving data. It also monitors the overall network performance that what is going on? If any misbehave action detects, it not only informs the surrounding nodes but also take some necessary action against those intruders. The ad hoc closed-key networks is comparatively more secure than the open ad hoc networks because closed-key networks have pre-define security policy for authentication and encryption but open ad hoc networks are free for any node to come in and becomes the part of the ad hoc network with arbitrary topology.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

In this paper a distributed local-IDS has proposed. Section-2 of the paper consists on related work in security for ad hoc networks, section-3 has a MANETs tread model and in section-4 the proposed system are discussed with pros and cons. Section-5 have the concluding remarks of the paper.

II. RELATED WORK

The traditional security mechanisms are insuring by using the concept of key management. But key management becomes difficult in the presence of an active attacker node. A reasonable solution is Certification Authority (CA) [3]. CA has a public and private key pairs. The public key of the CA is known to everyone and it makes a certificate of having the public key of each node sign by its private key [4]. This approach is valid with a massive overhead in the network because of dynamically changing topology of MANETs and every times verification of each valid node. Another issue is, if the CA node is being down, who is next CA? Multiple CAs is also recommended but still overhead created in the network. A distributed CAs concept also proposed but the problem remains the same and network experiences an extra overhead [5]. In fact, CA identifies each node have a valid certificate which prevent the spoofing and other malicious activities. But certificate verification requires a strong management system between CAs and surrounding nodes. But due to the limited resources of each node and unique characteristics of MANETs, it is implemented rarely and researchers want a feasible solution to reduce this overhead.

Symmetric key encryption is also used for authentication and authorization process for a node within the network. But network layer issues are encounter when such approach is used for ad hoc networks [6]. Localized certification is another approach which is based on public key infrastructure (PKI). The CAs and other nodes distribute secret shared updates with revocation list in such typical scenarios [7]. Another solution is Secure Routing Protocol (SRP), in which the correct routs are discovered from time to time so that compromised and re-played route are find out and must be discarded. Security associations exist between ends nodes because no intermediate nodes take participate in path discovery. The unique identifier number and authentication codes are used for correct rout discovery [8].

Many intrusion detection systems have also proposed. In [9], co-operative and distributed IDS for ad hoc networks have proposed which works on statistical anomaly based detection. In [13], based on Suburban Ad-hoc Network (SAHN) an intrusion detection system been proposed known as SAHN-IDS. SAHN-IDS useful for multi hop ad hoc network, where it detects misbehavior node by getting unfair share of transmission channel. It also detects anomalies in packet forwarding in effective and unique. The simulation results show the efficiency of the proposed scheme. In [14], a "Cross Layer Based Intrusion Detection System"(CIDS) has proposed for ad hoc networks. It detects intruders by analyzing the pattern of trace files. It communicates data securely from source to destination which increase network efficiency. Many other IDS for ad hoc network are proposed, but the principle is the same that all IDSs are design to protect the MANETs from outsider and insider attacks. The proposed local distributed-IDS are different in working mechanism from previous approaches. It is very effective in those situations where malicious code plays an important role in inside and outside network attacks.

III. THREAD MODEL

Ad hoc networks work in co-operation by dynamically changing topologies between mobile nodes. This property makes ad hoc network more vulnerable to active and passive attacks. Most of the attacks are meet in middle or denial of services (DOS) nature, which ranges from passive interfacing to active interfering. In MANETs, the DOS attack mostly launched due to the laptop nodes, which are rich in resources as compared to other nodes. In MANETs, DOS are launched in any layer, at physical layer the DOS attack is to constantly transmitting the signals which interferes the radio frequencies of the network. This can be done by one or more nodes. Continuous retransmitting jams the network and infected for desire purpose. Dos attacks are also launched on data link layer by violating the communication protocol (802.15.4 0r Zigbee) by continually transmitting messages in order to generate collisions. As such collisions would require retransmissions by the effected node it is possible to deplete the power of the node. In network layer, the DOS attack is launched on routing protocols [10]. In MANETs, one dedicated DOS attacks is Black hole router attack, the attacker node claim to be the shortest path node to surrounding nodes, getting information from surrounding and does not forwarded to the base station. Other type is resource exhaustion, in which the attacker node broad cast or uni-cast a massage (HELLO flood attack) to other nodes again and again, which results resources

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

consumption of the nodes resources like battery, CPU and memory [12]. A routing loop is another DOS attack, in which a loop is introduce in routing path, which results just circulate the information but not reach to the base station.

The meet in the middle (MIM) attack are also very obvious attack on MANETs. This attack is more easily launched due to the ad hoc nature of the network. In MIM, the existing resources of MANETs are utilized in such a way that they not only actively interferes the network traffic but also play a vital role as an eavesdropper. Many types of MIM attacks are discovered in MANETs, replication attacks one of them. In this attack node is captured, analyze, replicate and insert these replicas within the network. Another one is Sybil Attack, in which a single malicious node masquerading with multiple identities. This single node can then have a serious impact on fault-tolerant schemes such as distributed storage, data aggregation and multi-path routing [10]. The network attack is another one; the attacker node partitions the connecting network into mini sub networks. These sub networks are not communicated although they are connected [11]. The malicious node can also corrupt the data or miss routed it. The base station (BS) play very important role because it is the central point of aggregate data, all decisions about network management are decide on the base station. So if base station is compromised, the whole network is compromised, that is why the base station is protected from every promising attack.

IV. PROPOSED SYSTEM

Many IDS for ad hoc network have proposed. Some of them have critical for certain scenarios. Some of them are used with collaboration of routing protocols. Here we propose distributive local-IDS for ad hoc networks. This local-ID may be used for low energy nodes like sensor nodes. Sensor nodes have limited resources with special design purpose. The proposed IDS can also used for more power full mobile nodes, having more resources. It is distributive because each node in the network analyze the data individually and independently by smart agents and therefore each node have work as an IDS agent dispersed into the entire network. It is local because each node checks data/network behavior locally. And it is co-operative because it informs other nodes as well as base station. The base station then responsible for overall network performance and with the co-operation of other nodes it takes some necessary action against such hateful activity.

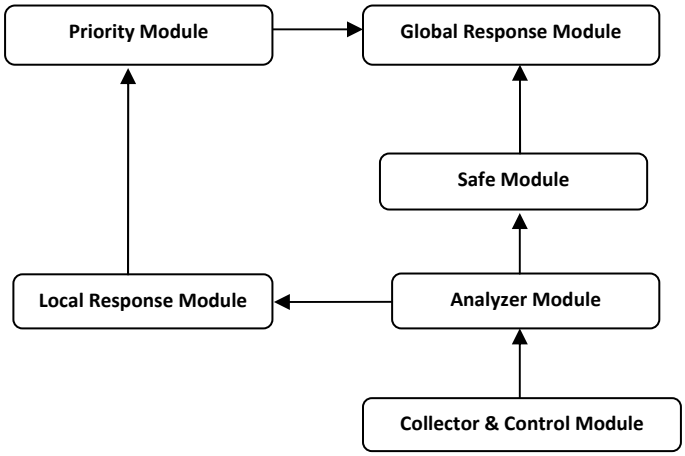


Fig.1 System Model of Local-IDS within a node

First the data is collected and then analyzed for intruders. After analysis an appropriate action is taken. Each node has their own local IDS agent for checking the received data. These agents have some previous signature or pre-define profile. When data is entered into these agents, the node first analyzes the receiving data. It analyzes data by comparing it for normal and abnormal activities with the threshold value of the pre-define profile. If some activity been detected as malicious, it must inform the base station or cluster head (CH) for further analysis. On the basis of investigation the base station or CH tacks an appropriate action. The targeted node may also inform the surrounding nodes, to aware of such falsified malicious data. The local IDS agent must be program in such way that it must detect normal and abnormal activities. The smart agent works on Markov process. Each node in the network updates its profiles/signature according to the base station commands. When base station receives the data having a complaint massage from the node, the base station first analyze the same abnormal behavior/malicious data. The base station informs rest of the cluster heads in that particular area and also informs other base station for this abnormal activity/malicious data. The base station now watches the overall network behavior and also waits the updates coming from other cluster heads as well as from other base stations. All these activities help the base station for checking the performance of the network. The base station sends updates to network nodes using Markov process. The last node in the hierarchy receives the difference of all of the nodes from base station to the last node. The net difference between two profiles/signatures is the signature updates.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

V. SYSTEM MODEL

The proposed system model consists of many parts. The main parts of the Local-IDS agent are shown in figure.1. First data is collected by collection and control module. It is "collector" because it collects data from other nodes. It is "controller" in a sense that it controls all the activities of the local IDS agent. Collected data then moved to analyzer module for analysis. The analyzer actually decided the working criteria. This part of the system depends upon the system design. Either works on protocol analysis (algorithm), pre-define profile or pre-define signature. The analyzer module is actually the key place where the base station maintains the pre-define signature or profile for each node. The updates from the base station to IDS agents are come through Markov process. If analyzer module is tightly design then it increases the false positive rate, which collected erroneous as well as correct data. But the analyzer module must also decreases the false negative, in which erroneous data is also marked as correct data. After analyzer, the data is either pass to the safe module or emergency module. Data in the safe module show normal data having no abnormal code. Safe module sends data to global response module (GRM) for sending base station on normal basis. Safe module plays an important role in data forwarding when priority being assign. The emergency module is also known as Local Response Module (LRM). If data is passing to local response module, it means the analyzer find something wrong in the data/system behavior. Consequently LRM send an alarm massage to surrounding nodes that all nodes should warn about such thread. The data then pass into priority module. Where priorities are assigned to those packets and send it to GRM. The GRM send that suspected data to the base station for further analysis. The base station then further analyze these packets and send massages to other base station and cluster heads. The base station also sends important messages to those nodes that sense the thread for first time in the network. The controllers of the IDS at each node receive those massages and responds accordingly. The base station checks the overall data flow, over all behavior of the network and receive massages from other base station as well. The base station then follow a procedure how to tackle the intruders and how mange the overall network. The base station communicates to leaf nodes by following the same route from base station to leaf node. The safe module is programmed in such a way to direct traffic from leaf node to base station and also from base station to leaf nodes. The intermediate nodes become as forwarding nodes that only forward the messages.

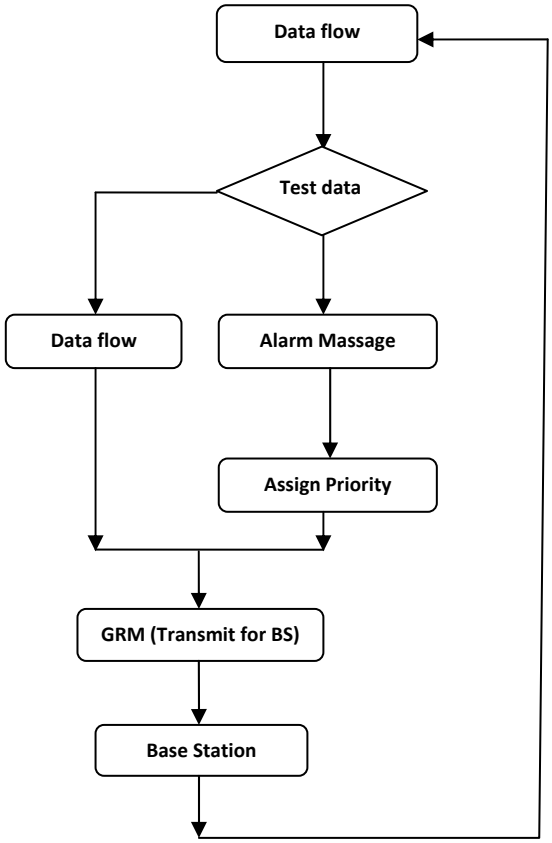


Fig.2 System Flow diagram.

The distributed IDS is actually the smart agents based IDS. The data is collected locally by these smart agents. If something find abnormal by comparing the profiles or signature. Then it sends those data on priority bases to base station and also informs the surrounding nodes about those malicious data. The base station is now monitors the overall network performance by analyzing the behavior of the nodes. For example if out of five hundred nodes two hundred are suddenly down or some existing paths are suddenly change. Then the base station look for those abnormal behavior and respond like a typical intrusion prevention system. It saves further network damage by responding on time to the leaf nodes. The base station is actually tells the controller of the agents what to do? How to do? And when to do? If the base station finds some malicious activity continuously acting on surrounding nodes (like in DOS attack), the base station sends message to controller that do not collect data until next commands. The base station also

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

can tells the nodes that this type data are not send to base station comparing to some signature. The base station tells the nodes to collect the data by sending a massage having one for collection and zero for dropping the data. The base station sends updated signature to the agents for comparison by using Markov process. In real situations the base station may be far away from sensing node. And the data is send through other nodes from leaf nodes to base station. For that case the data is not check by each node if some priority being assign to it. The priority assigned values is send first because it is important. An algorithm must maintain how to assign priority and how to send such packets before any data send. In fig.2, the system flow chart shows the overall structure of Local-IDS related to base station.

VI. SIMULATIONS AND FUTURE WORK

Consider a network having many nodes, each of them having an intrusion detection system (smart agents). These local-IDS are capable of checking the incoming packets to the MANETs. Consider the following simulation parameters. A network consisting of MANETs nodes having communication range from 150 t0 200 meters, covering an area of 600 by 600 square meters.

Topology shape	600 meter *600 meter
Radio Range of each node	200 meters
Node moments	Random/Zigzag
Base Station Moment/Static	Random/Zigzag
Topological Model	Multi hop planner/hierarchical
Maximum speed of a node	3-5 meters/second
Transmission Capacity	1.5 Mbps
Set Node count	15
Total flows	10-15
Average transmission per flow	2 packets per second
Testing execution time	40 seconds

Tables.1 Simulation Parameters

MANETs nodes can move any direction, the base station also randomly move. Maximum speed of each node 5 meter per second but it can also move with less velocity. Transmission capacity of each node is 1.5 Mbps, with initial set count of 20. Total flows in the network when initially test is 10. Testing execution time is 50 seconds, and average transmission flow of the network is 2 packets per second.

Each MANETs node updates it's pre-define signature/profile by using the Markov process. Markov process shows the difference of two events/variables. For example

$$A \rightarrow B \rightarrow C \rightarrow D \rightarrow E \rightarrow S^2$$

The value of (S^2) is the difference of all the previous events. Therefore, (C) shows the difference of (A) and (B), (D) have the difference of (C) and (B), (E) consists of difference of (D) and (C) and so on. So the (S^2) have the value which is different from all previous events but depends upon the values of (E) and (D). The same back tracking is true for other values in the hierarchy. In the following equations the difference shows at base station, the difference of all the nodes from leaf to base station. The nodes automatically updates it's signatures by using this Markov process.

$$\sum(x) = x1 + x2 + x3 + \cdots + xn$$

$$p(x1, x2, x3, x4, \ldots\ldots, xn) = \sum p(x1, x2), p(x2, x3)p(x3, x4) \ldots\ldots p(xn-1, xn)$$

In other words the current threshold of the leaf node is depends upon the previous state of the node or the threshold of the above node in the hierarchy. The following topology explains the process in brief.

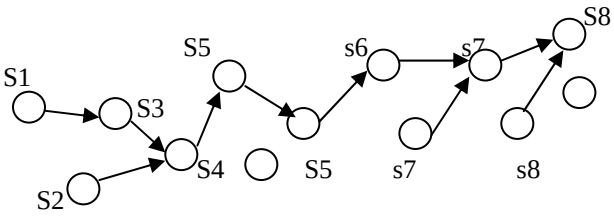


Fig.3 Ad hoc topology

In the above topology, S3 gets updates from S1 and S4 gets updates from S2 and S3 and so on up to S8 which near to the base station. It gets updates from base station. The base station also sends messages in the same way as receive messages. When analyzer node detects data as malicious, it assigns a priority to those packets. For example s1 detected such packets, then other nodes s2, s3, s4, s5, s6 do not check it, it just passed those packets to base station as quick as possible. The base station further analyzes the data and sends a massage to the cluster heads. As denial of service (DOS) attack is so common in MANETs. The local-IDS prevents such attacks by analyzing packets in term pre-define

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

profiles/signature and also monitoring the overall performance of the network at base station.

VII. CONCLUSION

Instead of proactive security mechanism some reactive security mechanism are required for MANETs, because the ad hoc nature of the network. In this paper we proposed Local-IDS, work locally in co-operative manner, locally analyzed the data/network behavior, if something is going in wrong direction, it not only inform local nodes but also inform the base station for further analysis. The distributed nature of local-IDS not only secures the ad hoc networks but also helps in that environment where no central management is ensuring like MANETs.

ACKNOWLEDGMENT

We are very thankful to Almighty Allah; whose grace and blessed mercy enabled us to complete this work with full devotion and legitimacy. We are grateful to Dr. Ata ul Aziz Ikram, Associate Professor & Head of the Department, Department of Computing & Technology, Iqra University Islamabad, for their invaluable support and guidance throughout this research work.

We also want to thank our friends and family for their encouragement; without whose support we could not have lived through this dream of ours.

VIII. REFERENCE

[1] Poly Sen, Nabendu Chaki, Rituparna Chaki "HIDS: Honesty-rate Based Collaborative Intrusion Detection System for Mobile Ad-Hoc Networks".
[2] "Cooperative Routing in Mobile Ad-hoc Networks: Current Efforts Against Malice and Selfishness." By Sonja Buchegger, Jean-Yves Le Boudec .
[3] M. Gasser, A. Goldstein, C. Kaufman, B. Lampson, "The Digital Distributed Systems Security Architecture," 12th National Computer Security Conference.
[4] Wensheng Zhang, R. Rao, Guohong Cao, GeorgeKesidis "SECURE ROUTING IN ADHOC NETWORKS AND A RELATED INTRUSION DETECTION PROBLEM".
[5] L. Zhouand Z. Haas, "Securing Ad Hoc Networks," IEEE Net-work
[6] Frank Stajano and Ross Anderson. "The Resurrecting Duckling." Lecture Notes in Computer Science, Springer-Verlag, 1999.
[7] Jiejun Kong, Petros Zerfos, Haiyun Luo, Songwu Lu, Lixia Zhang. "Providing Robust and Ubiquitous Security Support for Mobile Ad-Hoc Networks." In International Conference on Network Protocols (ICNP), pages 251–260, 2001
[8] Panagiotis Papadimitratos and Zygmunt J. Haas. "Secure Routing for Mobile Ad Hoc Networks" In SCS Communication Networks and Distributed Systems Modeling and Simulation Conference. (CNDS 2002), San Antonio, TX, January 2002
[9] Yongguang Zhang and Wenke Le " Intrusion Detection in Wireless Ad-Hoc Networks" In Proceedings of MOBICOM 2000
[10] Michael Healy, Thomas Newe, Elfed Lewis "Security for Wireless Sensor Networks: A Review" Optical Fibre Sensors Research Centre, Department of Electronic and Computer Engineering, University of Limerick, Limerick, Ireland.(2009).
[11] Yi-an Huang, Wenke Lee. " A Cooperative Intrusion Detection System for Ad Hoc Networks ".
[12] Ernesto Jiménez Caballero, "Vulnerabilities of Intrusion Detection Systems in Mobile Ad-hoc Networks-The routing problem".
[13] O. Kachirski and R. Guha, Intrusion Detection Using Mobile Agents in Wireless Ad Hoc Networks, Knowledge, July, 2002.
[14] Muhammad Mahmudul Islam, Ronald Pose and Carlo Kopp. "An Intrusion Detection System for Suburban Ad hoc Networks"

AUTHORS PROFILE

Muhammad Nawaz Khan is lecturer in Computer Science in Govt. College of Management Science. In 2008, he received Silver Medal in B.S. (Hons) degree in Computer Science from University of Malakand, K.P.K. Pakistan. He partially completed MS in Computer Communication Security at School of Electrical Engineering & Computer Science NUST Islamabad, Pakistan. In 2010, he worked as a Research Assistant in a project on "Distributed Computing" supported by Higher Education Commission of Pakistan. Currently he is working as Research Assistant at Shaheed Zulfikar Ali Butto Institute of Science & Technology Islamabad. His research is focused on Computer Information Security especially Computer Communication Security. He has also showed keen interest in Ad-hoc networks (MANETs, VANETs), wireless communications security and security related issues in distributed computing. He intended to proceed his studies(PhD) in any of the above mentioned fields.

Ishtiaq Wahid received his B.S. degree in information technology from University of Malakand at Chakdara, Dir lower, KPK, Pakistan, in 2007; the M.S. degree in Computer Science from Iqra University Islamabad Pakistan in 2009. He is currently pursuing the Ph.D. degree with Department of Computing & Technology Iqra University Islamabad Pakistan. In 2010, he joined in University of Malakand as a lecturer. Since 2010, he has been a lecturer with this Institute. His current research interests include Ad-hoc networks, wireless communications, and virtual reality environment.

Muhammad Ilyas Khatak received his B.S. (Hons) degree in information technology from University of Malakand at Chakdara, Dir lower, KPK, Pakistan, in 2009. Currently he is doing MS in Computer Science major in Information Security Management, from Shaheed Zulfikar Ali Butto Institute of Science & Technology (SZABIST) Islamabad, Pakistan. His research interests include Information Security including Ad-hoc network security, wireless communication security, hand over in ad hoc networks and forensic analysis.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

Image Retrieval Using Histogram Based Bins of Pixel Counts and Average of Intensities

H. B. Kekre
Sr. Professor
Department of Computer Engineering,
NMIMS University,
Mumbai, Vileparle, India
hbkekre@yahoo.com

Kavita Sonawane
Ph. D. Research Scholar,
Department of Computer Engineering
NMIMS University,
Mumbai, Vileparle, India
kavitavinaysonawane@gmail.com

***Abstract*—This In this paper we are introducing a novel technique to extract the feature vectors using color contents of the image. These features are nothing but the grouping of similar intensity levels in to bins into three forms. One of its form includes count of number of pixels, and other two are based on bins average intensity levels and the average of average intensities of R,G and B planes of image having some similarity amongst them. These Bins formation is based on the histograms of the R, G and B planes of the image. In this work each image separated into R, G and B planes. Obtain the histogram for each plane which is partitioned into two, three and four parts such that each part will have equal pixel intensity levels. As the 3 histograms are partitioned into 2, 3and 4 parts we could form 8, 27 and 64 bins out of it. We have considered three ways to represent the features of the image. First thing we taken into consideration is the count of the number of pixels in the particular bin. Second thing considered is calculate the average of the R, G and B intensities of the pixels in the particular bin and third form is based on average distribution of the total number of pixels with the average R, G, B intensities in all bins. Further some variations are made while selecting these bins in the process where query and database images will be compared. To compare these bins Euclidean distance and Absolute distance are used as similarity measures. First set of 100 images having less distances between their respective bins which are sorted into ascending order will be selected in the final retrieval set. Performance of the system is evaluated using the plots obtained in the form of cross over points of precision and recall parameters in terms of percentage retrieval for only out of first 100 images retrieved based on the minimum distance. Experimental results are obtained for augmented Wang database of 1000 bmp images from 10 different categories which includes Flowers, Sunset, Mountain, Building, Bus, Dinosaur, Elephant, Barbie, Mickey and Horse images. We have taken 10 randomly selected sample query images from each of the 10 classes. Results obtained for 100 queries are used in the discussion.**

Keywords-component; Histogram, Bins approach, Image retrieval, CBIR, Euclidean distance, Absolute distance.

I. Introduction (Heading 1)

This paper describes the new technique for Content Based Image Retrieval based on the spatial domain data of the image. CBIR systems are based on the use of spatial domain or frequency domain information. Many CBIR approaches uses local and global information such as color, texture, shape, edges, histograms, histogram bins etc to represent the feature vectors of the images [1], [2], [3], [4], [5]. Color is the most widely used visual feature which is independent of the image size and orientation. Many researchers have used color histograms as the color feature representation of the image for image retrieval. Most of these techniques are using global or local histograms of images, some are using equalized histogram bins, some are using local bins formation method using histograms of multiple image blocks [6], [7], [8], [9]. Main idea used in this paper is instead of changing the intensity distribution of the original image by taking the equalized histogram [10], [11]; we are using the original histograms of the image as it is. We are separating the image into R, G and B planes; obtain the histogram for each plane separately which is partitioned into two parts having equal pixel intensities. By taking R, G and B value of each pixel intensity of an image we are checking in which of the two parts of R, G, B histograms it falls respectively and then the bin for that pixel will be finalized where it will be counted [12]. Second thing we are taking into account is the intensities of the pixels in each of the 8 bins and new set of 8 bins is obtained in which each bin has the count of average of R, G, B intensity values of each pixel in that bin. A little variation is made in second types of bins is that we are taking average of average R, G, B values of all pixels in the respective bin count and a third set of bins holding average of average is formed. After analyzing the results of 8 bins, we have increased the no of bins from 8 to 27 and 64 by dividing the histogram of each plane into 3 and 4 parts respectively. Once the bins formation is done comparison process is performed to obtain the results and evaluate the system performance. Comparison of query and database images requires similarity measure. It is significant factor which quantifies the resemblance in database image and query image [13],[14]. Depending on the type of features, the formulation of the similarity measure varies greatly The different types of distances which are used by many typical CBIR systems are Mahalanobis distance [15], intersection distance [16], the Earth mover's distance (EMD), Euclidian distance [15], [17], and Absolute distance [19]. In this paper we are focusing on Euclidean distance and absolute distance as similarity measures, using this we are calculating the distance between the query and 1000 database image feature vectors. These distances are then sorted in ascending

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

order from minimum to maximum Out of these 1000 sorted distances images with respect to create these components, first 100 distances in ascending order are selected as images retrieved as there are 100 images of each class in the database [18]. Number of relevant images in these 100 images gives us the precision and recall cross over point (PRCP), which is the performance evaluation parameter of the system.

This paper is organized as follows: Section 2 will discuss the algorithmic view of the CBIR system based on 8, 27 and 64 bins using histogram plots. Section 3 describes the Role of the similarity measures in the CBIR system. Section 4.highlights the experimental results obtained along with the analysis. Finally section 5 summarizes the work done along with their comparative study.

II. ALGORITHMIC VIEW OF BINS FORMATION

A. *Feature Extraction and Formation of Feature Databases*

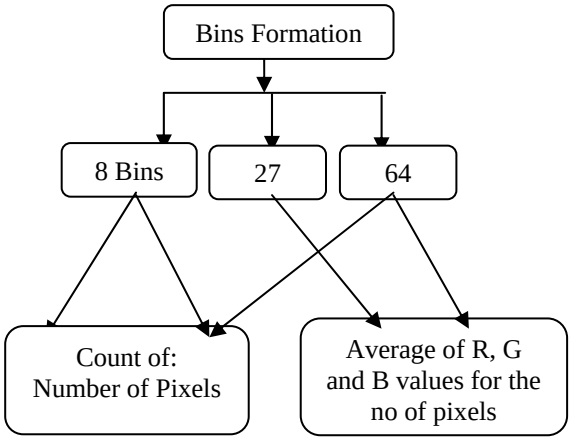


Figure 1. Feature vector Database Formation

Bins Formation Process: 8 Bins

Step1. Spilt the image into R, G and B planes.
Step2. Obtain the histogram for each plane.
Step3. Divide each histogram into 2 parts and assign a unique flag to each part.
*Step4***.** To extract the color feature of the image, pick up the original image pixel and check its R, G and B values find out in the histogram that in which range these values exactly falls, based on it assign the unique flags to the r, g and b values of that pixel with respect to the partition of the histogram it belongs.
Step5. Count of pixels in the bin: Based on the flags assigned to each pixel with respect to the R, G B values and 2 partitions (e. g. 0 and 1) of the histogram we can have 8 combinations from 000 to 111 which are the total 8 bins".

B. *Formation of Extended Bins 27 and Bins 64*

Formation of 27 and 64 bins feature vector database is extended version of the 8 bins feature extraction process. Here for 27 bins only difference is in step3 of the above algorithm, here to get 27 and 64 bins we are partitioning the histograms into 3 and 4 parts respectively which are named as 0, 1, 2 for 27 bins and 0, 1, 2, 3 for 64 bins approach. As explained in step 4 to 5 here also same process is applied and 3 bit flags are assigned to each pixel of the image for which the feature vector is being extracted. For 3 partitions the 3 flag bits (either of 0, 1 and 2) can have 27 combinations and for 4 partitions the 3 flag bits (either of 0, 1, 2 and 3) can have 64 combinations, these are the addresses of the 27 and 64 bins respectively. Based on this process two feature databases of feature vector size 27 and 64 holding the count of no of pixels according to the r, g, and b intensity values are obtained as Bins27_database and Bins64_database respectively.

C. *Variations to Obtain Multiple Feature Databases*

As shown in Figure.1 Three different databases for 8, 27 and 64 bins can further have 2 different sets of feature vectors named "Count of no of pixels", "Average of R, G and B values for all pixels in a Bin" which are simply obtained by modifying the process of extracting the feature vectors ; instead of just taking the count of pixels we have considered the significance of actual intensity levels of each pixel in each of the 8, 27 or 64 bins and taken the average values of them.

III. APPLICATION OF SIMILARITY MEASURE

Many similarity measures used in different CBIR systems are studied [21], [22], [23], [24], [25]. We have used Euclidean distance given in equation (1) and absolute distance in equation (2) as similarity measures in our work to produce the retrieval results. Once the query image is accepted by the system it will calculate the Euclidean distance as well as Absolute distance between the query image feature vector and database image feature vectors. In our system database size is 1000 images, so we obtained two sets of results one based on each similarity measure. When query image will be compared with 1000 database images which generate 1000 Euclidean distances and 1000 Absolute distances. These are then sorted in ascending order to select the images having minimum distance for the final retrieval.

Euclidean Distance :

$$D_{QI} = \sqrt{\sum_{i=1}^{n} \left|\left(FQ_i - FI_i\right)\right|^2} \quad (1)$$

Absolute Distance :

$$D_{QI} = \sum_{1}^{n} \left|\left(FQ_I - FI_I\right)\right| \quad (2)$$

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

Final Retrieval Process

Images having less distance are to be selected in the final set. For this we kept one simple criterion that we are taking first minimum 100 distances from the sorted list and corresponding images of those distances only taken into the final retrieval set. Same process is applied for all the features databases using both similarity measures

IV. Experimental Results And Discussions

A. Database and Query Images

Experimental set up for this work uses 1000 BMP images includes 10 different classes where each class has 100 images within it. The classes we have used are Flower, Sunset, Mountain, Building, Bus, Dinosaur, Elephant, Barbie, Mickey and Horse images. Feature vectors for all these images are calculated in advance using different methods described above in section 2 and multiple feature databases are obtained.

Query is given as example image to this system. Once the query enters into the system feature vectors using all different ways will be extracted and will be compared with the respective feature vector databases by calculating the Euclidean distance and Absolute distance between them. Selection of query images is from the database itself; it includes 10 images from each class means total 100 images are selected to be given as query to the system for all the approaches based on variations in bins formation to test and evaluate their performance. Sample Images from the database is shown in Figure 2.

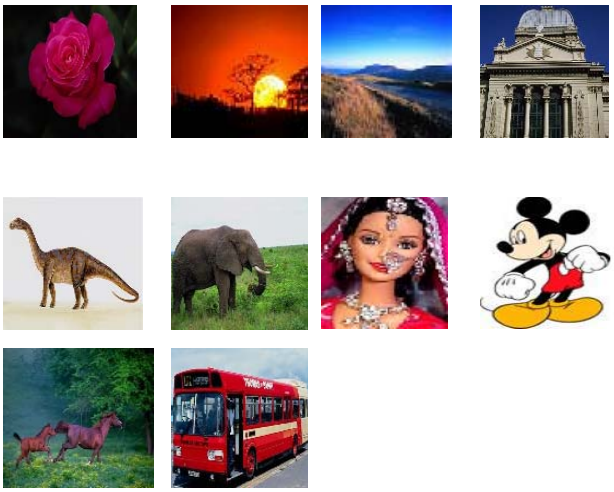


Figure 2. Sample Database Images from 10 Different Classes

(Database is of Total 1000 bmp images from above 10 classes, includes 100 from each class

B. Results, Observations and Comparison

Results using 100 queries are obtained for 3 approaches based on formation of bins, that are 8 bins, 27 bins and 64 bins where each approach includes the 2 variations while extracting the pixel's color information to form the feature vector which are classified as 'Count of Number of pixels' and 'Single average' that is average intensities of the number of pixels in each bin. Results obtained are segregated in three tables as 8 bins, 27 bins, and 64 bins. First column of each table is indicating the query image classes used for the experimentation. Remaining two columns are showing the total retrieval results obtained for Count of pixels and Single average approaches with respect to both the similarity measures that are Euclidean distance (ED) and Absolute distance (AD). Percentage retrieval is shown in Chart 1, 2 and 3 for 8, 27 and 64 bins respectively. Since there are 100 images of each class in the database percentage retrieval will be a cross over point of precision and recall [26].

In Table 1 we can see the total and average of retrieval of 10 queries from each of the 10 classes. In all the three results, results based on just the count of pixels are poor as compare to the other approaches. Results obtained for Single_Average are far better than 'Count of Number of Pixels'. We can note down the two sets of results are obtained for each approach; one is Euclidean distance and other is for Absolute distance named as ED and AD respectively. When we observe these results of ED and AD, we found that AD is giving very good performance as a similarity measure in both the approaches. Chart1 is showing the percentage retrieval where Single average proving its best for the class flower as it shows the highest retrieval that is almost 55%. After observing the results obtained for 8 bins we thought of extending these bins to 27 which are formed by dividing the histogram of each plane into 3 parts instead of 2 parts as in case of 8 bins.

TABLE I. Results for 8 Bins as Feature Vector

Query Images	**Count Of No of Pixels Total Retrieval**		**Single Average Total Retrieval**	
	ED	**AD**	**ED**	**AD**
Flower	246	253	480	547
Sunset	503	504	458	460
Mountain	161	170	236	252
Building	171	168	219	240
Bus	404	413	455	481
Dinosaur	216	234	375	342
Elephant	187	180	303	301
Barbie	165	173	289	273
Mickey	277	286	492	475
Horse	374	369	463	468
Average of 100 queries	2704	2750	3770	3839

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

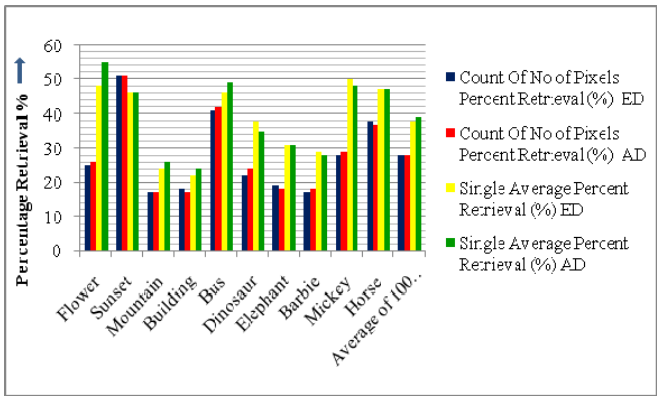


Chart 1. Results for 8 Bins as feature vector

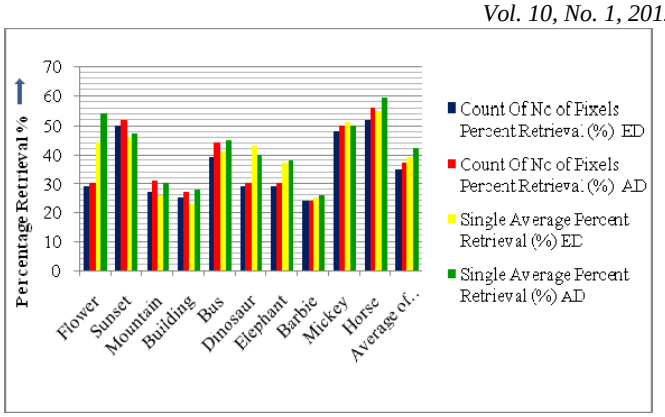


Chart 2. Results for 27 Bins as feature vector

Results obtained are shown in Table 2 and Chart2. Here noticeable positive change is obtained in the total retrieval of 'Count of No. of Pixels' approach. Single_Average' is also performing well as compare to the results of 8 bins.

Here also AD is giving very good retrieval results as compared to ED in all the cases. In Chart2 we can see that for the Horse class we got the highest percentage of retrieval that is around 59%.

This improvement in the results triggered us to further extend these bins from 27 to 64 by dividing the histogram into 4 parts which is generating the 64 bins. When we compared the results of 64 bins with the results for 8 and 27 bins, the performance is decreasing for Single_Average' and in case of 'Count of No. of Pixels' it is improved as compared to 8 bins but is little poor as compared to 27 bins. In this case when we observe Chart 3 it shows that both the approaches with absolute distance are giving best results for class horse, which is around 62%.

TABLE II. RESULTS FOR 27 BINS AS FEATURE VECTOR

Query Images	Count Of No of Pixels Total Retrieval		Single Average Total Retrieval	
	ED	AD	ED	AD
Flower	287	299	433	538
Sunset	496	515	451	461
Mountain	264	310	255	292
Building	243	268	226	277
Bus	383	435	407	447
Dinosaur	285	294	423	393
Elephant	284	293	368	373
Barbie	231	239	250	256
Mickey	480	494	502	497
Horse	520	553	543	583
Average of 100 queries	3473	3700	3858	4117

TABLE III. RESULTS FOR 64 BINS AS FEATURE VECTOR

Query Images	Count Of No of Pixels Total Retrieval		Single Average Total Retrieval	
	ED	AD	ED	AD
Flower	291	328	438	550
Sunset	460	480	394	420
Mountain	260	327	281	300
Building	249	280	242	300
Bus	322	454	342	400
Dinosaur	216	308	281	338
Elephant	284	312	287	308
Barbie	225	230	225	226
Mickey	487	521	497	490
Horse	601	612	513	615
Average of 100 queries	3395	3852	3500	3947

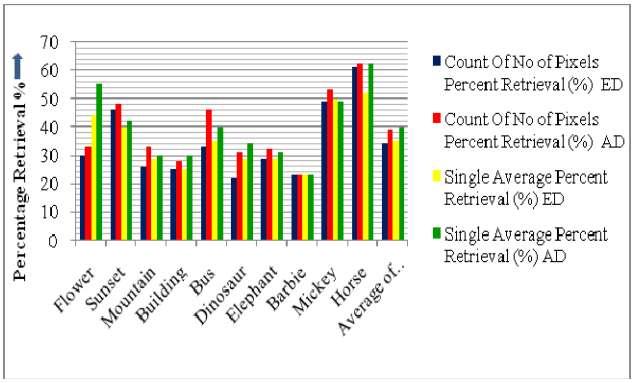


Chart 3. Results for 64 Bins as feature vector

When we compare overall results just on the percentage retrieval of all the classes taken into consideration, we can delineate that both approaches of feature vectors of size 27 bins are performing better as compare to 8 and 64 bins. Within that AD is giving far better results as compare to ED for all three results sets of 27 bins.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

All the charts are highlighting that among the results in all types of bins; Single Average with AD is performing well in terms of percentage retrieval. Last data point plotted in all the charts that is Average of 100 queries, shows that Single average AD is having percentage retrieval of 39 % for 8 bins, 42 % for 27 bins and 40% for 64 bins in Charts 1, Chart 2 and Chart 3 respectively.

Sunset Query

Retrieval…

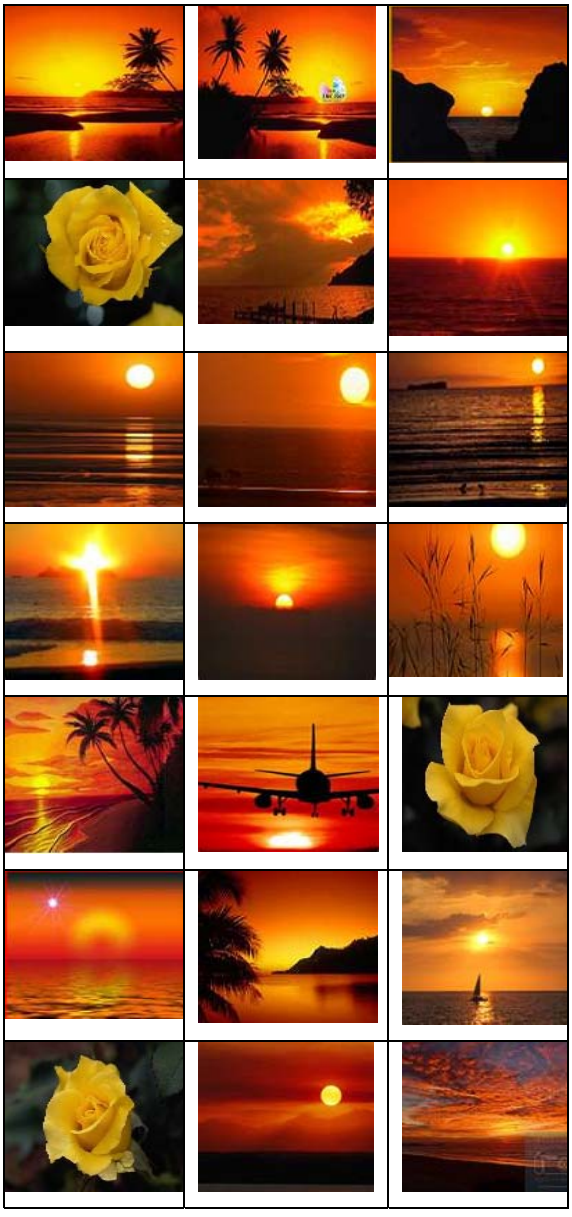


Figure 3. Sample Result of First 21 Images Retrieved

(63 Relevant images were retrieved in first 100 images)

Results shown in Figure 3 are the first 21 images retrieved for one of the randomly selected sunset query. It is observed that out of 21 images there are only three irrelevant images which happened to be flowers. This is good performance.

In all the approaches discussed above, feature vector extraction is mainly based on the color information. We have taken the separate histograms of the R, G, B planes of the image and while extracting the features we consider the R, G and B intensities of each pixel to see that which part of histogram it falls which actually determines the bin address of that pixel where it has to reside. This process is concentrating on the difference in the intensities that means mainly on color. Further analysis is done for these results with respect to the images, mainly their colors in the databases. This analysis is indicating that the 10 classes considered having 100 images each, are of different shapes and textures. With such a database, even though we have considered only color information in our approaches, we are getting very good retrieval result with less computational complexity.

V. Conclusion

In this work, all the approaches discussed above are based on the color information extraction in histogram based bins of count of number of pixels and their average intensities. Results are based on two measures of similarity that are Euclidean and Absolute distance mentioned in equation (1) and (2) respectively.

Results are obtained for two approaches that are, count of pixels and their average intensities for 3 different set of feature databases having 3 different sizes of feature vectors as 8 bins, 27 bins and 64 bins sets.

Among these results, if we compare them on the basis of bins-size, 27 bins approach is performing better as compared to other two.

When we compared the two approaches in all the bins that are: count of pixels and average intensities, we found that average intensities are producing promising results. This indicates that, instead of just taking the count of pixels, consider the intensities they have.

Results compare on the basis of similarity measures used, ED and AD as explained earlier, are suggesting that Absolute distance is giving very good results in all the cases and for all size of feature vectors. Same can be noticed in charts 1, 2 and 3 where green and red color bars are highlighting the results of absolute distance which are achieving good hight in the percentage retrieval.

References

[1] Darshak G. Thakore1, A. I. Trivedi, “Content based image retrieval techniques – Issues, analysis and the state of the art” www.rimtengg.com.

[2] Eva Gutsmiedl , “Content-Based Image Retrieval :Color Histograms”, , May 13th, 2004 URL of this document: http://www.fmi.uni-passau.de/~gutsmied/seminar/seminar.pdf

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

[3] Y. Rui, T. S. Huang and S. Chang, "Image Retrieval: Current Techniques, Promising Directions and Open Issues ", Journal of Visual Communication and Image Representation, vol. 10, pp. 39]62, March 1999.

[4] J. R. Smith and S.F. Chang, "Automated image retrieval using color and texture", Technical Report CU/CTR 40814, Columbia University, July 1995.

[5] J. Han and K. Ma, "Fuzzy Color Histogram and Its Use in Color Image Retrieval",IEEE Trans. On Image, Processing, vol. 11, zpp. 944 – 952, Aug. 2002.

[6] N.K.Kamila, ,Pradeep Kumar Mallick, Sasmita Parida B.Das, "Image Retrieval using Equalized Histogram Image Bins Moments" December 2010.

[7] Shengjiu Wang, A Robust CBIR Approach Using Local Color Histograms, Technical Report TR 01-03, Departement of computing science, University of Alberta, Canada. October 2001.

[8] A Vadivel , A K Majumdar, Shamik Sural , "Perceptually Smooth Histogram Generation from the HSV Color Space for Content Based Image Retrieval"

[9] M. J. Swain and D.H. Ballard. "Color indexing". In International Journal of Computer Vision, Vol. 7(1), pp 11-32, 199.

[10] Jeff Berens., "Image Indexing using Compressed Colour Histograms", Thesis submitted for the Degree of Doctor of Philosophy in the School of information Systems, University of East Anglia, Norwich.

[11] Greg Pass and Ramin Zabih. "Comparing Images Using Joint Histograms". ACM Journal of multimedia Systems, Vol. 7(3), pp. 234-240, May 1999.

[12] Guoping Qiu "Color Image Indexing Using BTC" IEEE Transactions On Image Processing, Vol. 12, No. 1, January 2003.

[13] C. Schmid and r. Mohr, "local grayvalue invariants for image retrieval," IEEE trans. Pattern anal. Mach. Intell., vol. 19, no. 5, pp. 530–535, may 1997.

[14] S. Santini and r. Jain, "similarity measures," IEEE trans. Pattern anal.mach. Intell., vol. 21, no. 9, pp. 871–883, sep. 1999.

[15] Y. Rubner, l. J. Guibas, and c. Tomasi, "The Earth mover's distance, multi-dimensional scaling, and color-based image retrieval," In proc.darpa image understanding workshop, may 1997, pp. 661–668.

[16] J. Hafner, h. S. Sawhney, w. Equitz, m. Flickner, and w. Niblack, "efficient color histogram indexing for quadratic form distance functions," IEEE trans. Pattern anal. Mach. Intell., vol. 17, no. 7, pp. 729–736, jul. 1995.

[17] Qasim Iqbal And J. K. Aggarwal, "Cires: A System For Content-Based Retrieval In Digital Image Libraries" Seventh International Conference On Control, Automation, Robotics And Vision (Icarcv'02), Dec 2002, Singapore.

[18] H. B. Kekre , Kavita Sonawane, "Query Based Image Retrieval Using kekre's, DCT and Hybrid wavelet Transform Over 1st and 2nd Moment" International Journal of Computer Applications (0975 – 8887), Volume 32– No.4, October 2011

[19] H.B.Kekre ,Dhirendra Mishra, "Sectorization of DCT-DST Plane for Column wise Transformed Color Images in CBIR" ICTSM-11, at MPSTME 25-27 February, 2011. Uploaded on Springer Link

[20] H. B. Kekre , Kavita Sonawane "Feature Extraction in Bins Using Global and Local thresholding of Images for CBIR" International Journal Of Computer Applications In Applications In Engineering, Technology And Sciences, ISSN: 0974-3596 | October '09 – March '10 | Volume 2 : Issue 2.

[21] Young-jun Song, Won-bae Park, Dong-woo Kim, and Jae-hyeong Ahn, "Content-based image retrieval using new color histogram", Intelligent Signal Processing and Communication Systems, Proceedings of 2004 International Symposium on 18-19 Nov. 2004, pp. 609-611.

[22] J. Huang, S. R. Kumar, M. Mitra, W. J. Zhu and R. Zabih, "Image Indexing Using Color" Proc.IEEE Conf. on Computer Vision and Pattern Recognition.

[23] Remco C. Veltkamp, mirela tanase department of computing science, utrecht university, "content-based image retrieval systems:a survey" Revised and extended version of technical report uu-cs- 2000-34, october october 28, 2002.

[24] H. B. Kekre, Kavita Sonawane "Standard Deviation of Mean and Variance of Rows and Columns of Images for CBIR" WASET International Journal of Computer, Information and System Science and Engineering (IJCISSE), Volume 3, Number 1, pp.8-11, 2009

[25] Yixin chen, member IEEE, james z. Wang, member IEEE, and robert krovetz clue: "Cluster-Based Retrieval Of Images By Unsupervised Learning" IEEE Transactions On Image Processing, Vol.14, No. 8, August 2005.

[26] Dr. H. B. Kekre, Sudeep D. Thepade, Varun K. Banura, "Performance Comparison of Gradient Mask Texture Based Image Retrieval Techniques using Walsh, Haar and Kekre Transforms with Image Maps" International Journal of Computer Applications (IJCA), Special Issue July 2011. Selected as Editors Choice(Best Paper)

AUTHORS PROFILE

Dr. H. B. Kekre has received B.E. (Hons.) in Telecomm. Engg. from Jabalpur University in 1958,M.Tech (Industrial Electronics) from IIT Bombay in 1960, M.S. Engg. (Electrical Engg.) from University of Ottawa in 1965 and Ph.D. (System Identification) from IIT Bombay in 1970. He has worked Over 35 years as Faculty of Electrical Engineering and then HOD Computer Science and Engg. at IIT Bombay. For last 13 years worked as a Professor in Department of Computer Engg. at Thadomal Shahani Engineering College, Mumbai. He is currently Senior Professor working with Mukesh Patel School of Technology Management and Engineering, SVKM's NMIMS University, Vile Parle(w), Mumbai, INDIA. He has guided 17 Ph.D.s, 150 M.E./M.Tech Projects and several B.E./B.Tech Projects. His areas of interest are Digital Signal processing, Image Processing and Computer Networks. He has more than 350 papers in National / International Conferences / Journals to his credit. Recently twelve students working under his guidance have received best paper awards. Five of his students have been awarded Ph. D. of NMIMS University. Currently he is guiding eight Ph.D. students. He is member of ISTE and IETE.

Ms. Kavita V. Sonawane has received M.E (Computer Engineering) degree from Mumbai University in 2008, currently Pursuing Ph.D. from Mukesh Patel School of Technology, Management and Engg, SVKM's NMIMS University, Vile-Parle (w), Mumbai, INDIA. She has more than 8 years of experience in teaching. Currently working as a Assistant professor in Department of Computer Engineering at St. Francis Institute of Technology Mumbai. Her area of interest is Image Processing, Data structures and Computer Architecture. She has 7 papers in National/ International conferences / Journals to her credit.She is member of ISTE.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

The Increase Of Network Lifetime By Implementing The Fuzzy Logic In Wireless Sensor Networks

Indrit Enesi
Department of Electronic and Telecommunication
Polytechnic University of Tirana
Tirana, Albania
ienesi @fti.edu.al

Elma Zanaj
Department of Electronic and Telecommunication
Polytechnic University of Tirana
Tirana, Albania
ezanaj@fti.edu.al

Abstract
Wireless Sensor Networks (WSNs) present a new generation of real-time embedded systems with limited computation, energy and memory resources. They are being used in a wide variety of applications where traditional networking infrastructure is practically infeasible. Appropriate cluster-head node election can drastically reduce the energy consumption enhancing so the network lifetime. In this paper, a fuzzy logic approach to cluster-head election is proposed based on three descriptors - energy, concentration and centrality of nodes. Simulation shows that depending upon network configuration, a substantial increase in network lifetime can be accomplished as compared to probabilistically selecting the nodes as cluster-heads using only local information.

Key words — Wireless Sensor Networks, Network-lifetime, Cluster-head, Fuzzy Logic.

1. INTRODUCTION

With the recent advances in Micro Electro-Mechanical Systems (MEMS) technology, low power digital circuitry and RF designs, WSNs are considered to be one of the potential emerging computing technologies, edging closer towards widespread feasibility [5]. Several useful and varied applications of WSNs include applications requiring information gathering in harsh, inhospitable environments, weather and climate monitoring, detection of chemical or biological agent threats, and healthcare monitoring. These applications demand the usage of various equipment including cameras, acoustic tools and sensors measuring different physical parameters [7]. The energy supply of the sensor nodes is one of the main constraints in the design of this type of network [6]. Since it is infeasible to replace batteries once WSNs are deployed, an important design issue in WSNs is to lessen the energy consumption with the use of energy conserving hardware, operating systems and communication protocols. The energy consumption can be reduced by allowing only some nodes to communicate with the base station. These nodes called cluster-heads collect the data sent by each node in that cluster compressing it and then transmitting the aggregated data to the base station [1]. Appropriate cluster-head selection can significantly reduce energy consumption and enhance the lifetime of the WSN. In this paper, a fuzzy logic approach to cluster-head election is proposed based on three descriptors - energy, concentration and centrality. Simulation shows that depending upon network configuration a substantial increase in network lifetime can be accomplished as compared to probabilistically selecting the nodes as cluster-heads using only local information. There are diverse applications of intelligent techniques in wireless networks [4]. Fuzzy logic control is capable of making real time decisions, even with incomplete information. We compare our approach to a previously proposed popular cluster-head selection technique called LEACH (Low Energy Adaptive Clustering Hierarchy) [1]. LEACH is based on a stochastic model and uses localized clustering. The nodes select themselves as cluster-heads without the base station processing. Other nodes in the vicinity join the closest cluster-heads and transmit data to them. Simulation results show that our approach increases the network lifetime considerably as compared to LEACH

2. RELATED WORK

A typical WSN architecture is shown in Figure 1. The nodes send data to the respective cluster-heads, which in turn compresses the aggregated data and transmits it to the base station.

Many proposals have been made to select cluster-heads. In the case of LEACH [1], to become a cluster- head, each node n chooses a random number between 0 and 1. If the number is less than the threshold T(n), the node becomes the cluster-head for the current round.

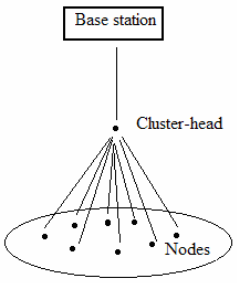


Fig. 1: WSN architecture

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

The threshold is set at:

If $n \in G$

$$T(n) = \frac{P}{1 - P(r \bmod \frac{1}{P})} \quad (1)$$

$T(n) = 0$ otherwise

where P is the cluster-head probability, r the number of the current round and G the set of nodes that have not been cluster-heads in the last 1/P rounds. Several disadvantages are there for selecting the cluster-head using only the local information in the nodes. Firstly, since each node probabilistically decides whether or not to become the cluster-head, there might be cases when two cluster-heads are selected in close vicinity of each other increasing the overall energy depleted in the network. Secondly, the number of cluster-head nodes generated is not fixed so in some rounds it may be more or less than the preferred value. Thirdly, the node selected can be located near the edges of the network; wherein the other nodes will expend more energy to transmit data to that cluster-head. Fourthly, each node has to calculate the threshold and generate the random numbers in each round, consuming CPU cycles. LEACH-C [2] uses a centralized algorithm and provides another approach to form clusters as well as selecting the cluster-heads using the simulated annealing technique. In [3] each node calculates its distance to the area centroid which will recommend nodes close to the area centroid and not the nodes that is central to a particular cluster, cluster centroid.

3. SYSTEM MODEL

In this paper the cluster-heads are elected by the base station in each round by calculating the chance each node has to become the cluster-head by considering three fuzzy descriptors, node concentration, energy level in each node and its centrality with respect to the entire cluster. In our opinion a central control algorithm in the base station will produce better cluster-heads since the base station has the global knowledge about the network. Moreover, base stations are many times more powerful than the sensor nodes, having sufficient memory, power and storage. In this approach energy is spent to transmit the location information of all the nodes to the base station (possibly using a GPS receiver). The operation of this fuzzy cluster-head election scheme is divided into two rounds each consisting of a setup and steady state phase similar to LEACH. During the setup phase the cluster-heads are determined by using fuzzy knowledge processing and then the cluster is organized. In the steady state phase the cluster-heads collect the aggregated data and performs signal processing functions to compress the data into a single signal. The energy expended during transmission and reception for a k bit message to a distance d between transmitter and receiver node is given by:

$$E_{Tx}(k, d) = E_{elec} * k + \varepsilon_{amp} * k * d^{\lambda} \quad (2)$$

$$E_{Rx}(k) = E_{elec} * k \quad (3)$$

where, λ is the path loss exponent and $\lambda \geq 2$.

3.1. Fuzzy Logic Control

The model of fuzzy logic control consists of a fuzzifier, fuzzy rules, fuzzy inference engine, and a defuzzifier. We have used the most commonly used fuzzy inference technique called Mamdani Method [8] due to its simplicity. The process is performed in four steps:

- Fuzzification of the input variables energy, concentration and centrality - taking the crisp inputs from each of these and determining the degree to which these inputs belong to each of the appropriate fuzzy sets.
- Rule evaluation - taking the fuzzified inputs, and applying them to the antecedents of the fuzzy rules. It is then applied to the consequent membership function (Table 1).
- Aggregation of the rule outputs - the process of unification of the outputs of all rules.
- Defuzzification - the input for the defuzzification process is the aggregate output fuzzy set chance and the output is a single crisp number.

During defuzzification, it finds the point where a vertical line would slice the aggregate set chance into two equal masses. In practice, the COG (Center of Gravity) is calculated and estimated over a sample of points on the aggregate output membership function, using the following formula:

$$COG = (\Sigma \mu_A(x) * x) / \Sigma \mu_{A(x)} \quad (4)$$

where, μ_A is the membership function of set A.

3.2. Expert Knowledge Representation

Expert knowledge is represented based on the following three descriptors:

- Node Energy - energy level available in each node, designated by the fuzzy variable energy,
- Node Concentration - number of nodes present in the vicinity, designated by the fuzzy variable concentration,
- Node Centrality - a value which classifies the nodes based on how central the node is to the cluster, designated by the fuzzy variable centrality.

To find the node centrality, the base station selects each node and calculates the sum of the squared distances of other nodes from the selected node. Since transmission energy is proportional to d^2 (2), the lower the value of the centrality, the lower the

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

amount of energy required by the other nodes to transmit the data through that node as cluster-head. The linguistic variables used to represent the node energy and node concentration, are divided into three levels: *low*, *medium* and *high,* respectively, and there are three levels to represent the node centrality: *close*, *adequate* and *far,* respectively. The outcome to represent the node cluster-head election chance was divided into seven levels: *very small*, *small*, *rather small*, *medium*, *rather large*, *large*, and *very large*.
The fuzzy rule base currently includes rules like the following: if the *energy* is *high* and the *concentration* is *high* and the *centrality* is *close* then the node's cluster-head election *chance* is *very large*. Thus we used 33 = 27 rules for the fuzzy rule base. We used triangle membership functions to represent the fuzzy sets *medium* and *adequate* and trapezoid membership functions to represent *low*, *high*, *close* and *far* fuzzy sets. The membership functions developed and their corresponding linguistic states are represented in Table 1 and Figures 2 through 5.

Table 1. Fuzzy rule base.

	energy	concentration	centrality	chance
1	low	low	close	small
2	low	low	adeq	small
3	low	low	far	vsmall
4	low	med	close	small
5	low	med	adeq	small
6	low	med	far	small
7	low	high	close	rsmall
8	low	high	adeq	small
9	low	high	far	vsmall
10	med	low	close	rlarge
11	med	low	adeq	med
12	med	low	far	small
13	med	med	close	large
14	med	med	adeq	med
15	med	med	far	rsmall
16	med	high	close	large
17	med	high	adeq	rlarge
18	med	high	far	rsmall
19	high	low	close	rlarge
20	high	low	adeq	med
21	high	low	far	rsmall
22	high	med	close	large
23	high	med	adeq	rlarge
24	high	med	far	med
25	high	high	close	vlarge
26	high	high	adeq	rlarge
27	high	high	far	med

Legend: adeq=adequate, med=medium, vsmall=very small, rsmall=rather small, vlarge=very large, rlarge=rather large.

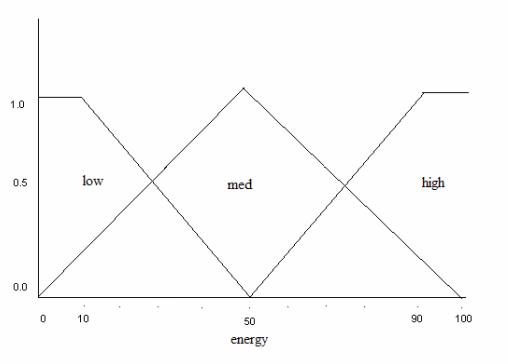


Fig. 2: Fuzzy set for fuzzy variable *energy*

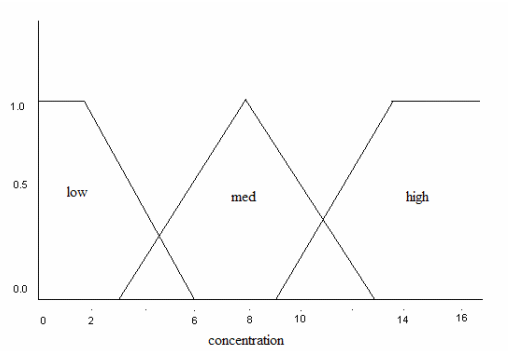


Fig.3: Fuzzy set for fuzzy variable *concentration*

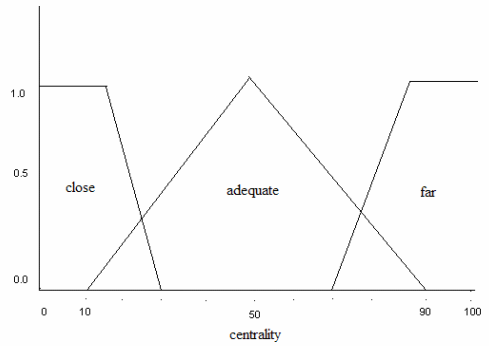


Fig. 4: Fuzzy set for fuzzy variable *centrality*

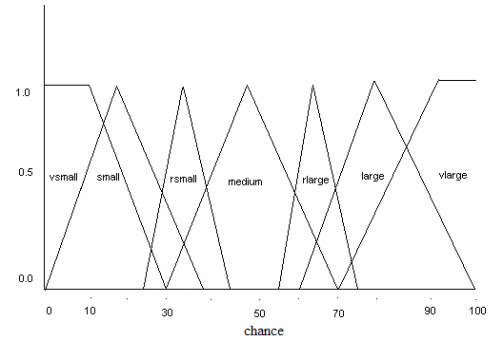


Fig. 5: Fuzzy set for fuzzy variable *chance*

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

4. RESULTS

To test and analyze the algorithm, experimental studies were performed. The simulator was programmed using Java Foundation Classes and the NRC fuzzy Java Expert System Shell (JESS) toolkit. We modelled the energy consumption in WSN as given in (2, 3). To define the lifetime of the sensor network we used the metric First Node Dies (FND) [9], meant to provide an estimate for the quality of the network.

4.1. Sample network 1

The reference network consists of 150 nodes randomly distributed over an area of 100x100 meters. The base station is located at 200, 50. In the first phase of the simulation each node has a random energy between 0 and 100. The base station computes the concentration for each node by calculating the number of other nodes within the area of 20x20 meters, with that node in the centre. The values are then fuzzified and passed to the fuzzy rule base for rule evaluation. After this, defuzzification gives the cluster-head election chance. Figure 7 shows the defuzzified output and the aggregate set chance for a specific node. The best nodes in terms of fuzzy overall, centrality and energy are shown in Fig. 6. Illustrating the results we can see that the best energy node has a very high centrality of 41 implying the overall energy spent by other nodes to transmit through node 62 will be high and hence a low cluster-head election chance. The best node 108 on the other hand has all the three descriptors suitable for being elected as the cluster-head with a maximum chance of 75 for the current scenario.

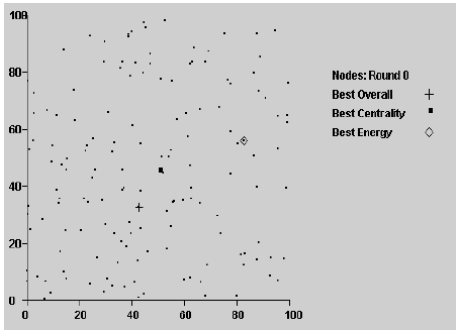


Fig. 6: Network cluster showing the best nodes

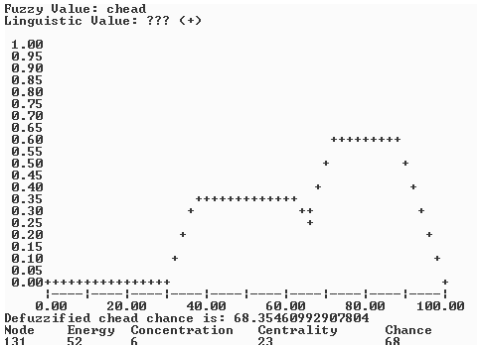


Fig. 7: Output fuzzy set for fuzzy variable *chance*

4.2. Sample network 2

In this case each node is supplied with energy of 1J at the beginning of the simulation. The *energy* fuzzy set is scaled accordingly, other parameters remaining unaltered. Each node transmits a 200 bit message, per round, to the elected cluster-head. The path loss exponent λ is set at 2 for intra-cluster communication and 2.5 for base station transmission. Cluster-head compresses the collected data to 5% of its original size. Figure 8 shows a snapshot of the simulation run for round number 44 with fuzzy elected cluster-head nodes. Figure 9 shows parameters for elected cluster- heads during two consecutive rounds 43 and 44. It takes about 2500 rounds for the FND in the network.

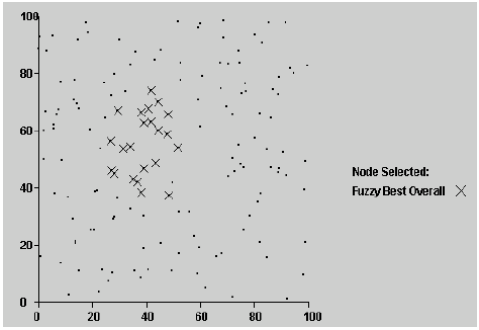


Fig. 8: Simulation in progress

```
Round: 43

Fuzzy Best node overall:
Node no: 1      CoOrdn: 43 48          Chance: 72
Energy: 90377   Conc:  8               Centrality: 21

Energy Spent at the Clusterhead 4790

                ---------------------

Round: 44

Fuzzy Best node overall:
Node no: 33     CoOrdn: 51 54          Chance: 71
Energy: 91512   Conc:  6               Centrality: 21

Energy Spent at the Clusterhead 4222
```

Fig. 9: Elected Cluster-heads for two consecutive rounds

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

5. CONCLUSION

This paper has discussed a novel approach for cluster-head election for WSNs. Cluster-heads were elected by the base station in each round by calculating the chance each node has to become the cluster-head using three fuzzy descriptors. Our approach is more suitable for electing cluster-heads for medium sized clusters. With this system model a substantial increase in the network lifetime is accomplished as compared to LEACH. By modifying the shape of each fuzzy set accurately, a further improvement in the network lifetime and energy consumption can be achieved. Since centrality, calculated on the basis of the sum of the squared distances of other nodes from the given node, is one of the descriptors for electing suitable cluster-head, a network with biased distribution of nodes can be tested in the future with further experiments.

6. REFERENCES

[1] W. Heinzelman, A. Chandrakasan and H. Balakrishnan, "Energy-efficient communication protocol for wireless microsensor networks," in Proc. of the 33rd AnnualHawaii International Conference on System Sciences (HICSS), Maui, HI, Jan. 2000, pp. 3005 – 3014.

[2] W. Heinzelman, A. Chandrakasan and H. Balakrishnan, "An application-specific protocol architecture for wireless microsensor networks," in IEEE Transactions on Wireless Communications, Oct. 2002, pp. 660 - 670.

[3] Q. Liang, "Clusterhead election for mobile ad hoc wireless network," in Proc. 14th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications, (PIMRC), Sept. 2003, pp. 1623 - 1628.

[4] S. Hammadi and C. Tahon, "Special issue on intelligent techniques in flexible manufacturing systems," in IEEE Transactions on Systems, Man and Cybernetics, May 2003, pp. 157 - 158.

[5] B. Warneke, M. Last, B. Liebowitz and K.S.J. Pister, "Smart Dust: communicating with a cubic-millimeter computer," IEEE Computer, Jan. 2001, pp. 44 - 51.

[6] E. Cayirci, "Data aggregation and dilution by modulus addressing in wireless sensor networks," IEEE Communications Letters, Aug. 2003, pp. 355 – 357.

[7] C. Chee-Yee and S.P. Kumar, "Sensor networks: evolution, opportunities, and challenges," in Proc of the IEEE, Aug. 2003, pp.1247 - 1256.

[8] M. Negnevitsky, Artificial intelligence: A guide to intelligent systems, Addison-Wesley, Reading, MA, 2001.

[9] M.J. Handy, M. Haase and D. Timmermann, "Low energy adaptive clustering hierarchy with deterministic cluster-head selection," in Proc. 4th International Workshop on Mobile and Wireless Communications Network, Sept. 2002, pp. 368 - 372.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

Mathematical Model for Component Selection in Embedded System Design

Ashutosh Gupta[#1], Chandan Maity[#2]

[#]*Embedded Systems Group,*
Centre for Development of Advanced Computing (C-DAC),
Noida, India
[1] ashutoshgupta@cdac.in
[2] chandanmaity@cdac.in

***Abstract*— Changes in embedded technologies and market dynamics have made traditional electronic parts selection and management practices inadequate. Component selection is a process designed to evaluate the electronic part, and facilitate informed decisions regarding its selection and future use. Embedded Designers face challenges when they are about to select the electronic component, for new design as it is difficult to compare the parts in terms of quantitative and qualitative terms in absence of any mathematical model. This paper proposes a new hybrid model which combines Linear Weightage and Analytic Hierarchy Process (AHP) Models linear weightage model to assist in the decision making activity and helps to select the best electronic component among a number of potential candidates. The final decision from this new model will help in better selection methodology for assisting embedded designers to make the right decision and select the most suitable component required for the design from the large pool of the components available in the market.**

Keywords - Mathematical Model, Component Selection, Embedded System Design, Linear Weightage Model, Analytic Hierarchy Process, Microcontroller

I. INTRODUCTION

The component selection and management methodology has been designed to aid in making risk informed decisions regarding the selection and use of electronic parts. The process aids in determining the acceptability of a component for an application, while considering factors such as functionality, performance, standardization, cost, availability, technology (new and aging), and logistics support.

Component selection is a process of selecting devices for the board design based on the various requirements like functional, electrical, mechanical, thermal, etc. Selection of a wrong component can create major problems in the functionality of the board. Hence, component selection is a very important aspect in the board design cycle. Component selection is a critical step, which will have lot of impact on rest of the project from the point of view of meeting functionality, performance, testing, manufacturing, confirming to standards and also to the schedule. In a typical product design cycle, component selection is required in the 3 following phases: Before a new design – new component selection; Component obsolescence – replacement with an updated version; Performance or feature enhancement – replacement with enhanced features.

Embedded Designers are often responsible for making purchasing decisions which is definitely a difficult task. There are many reasons which make the selection process a complex one, and the major are [1]:

- ✓ Component selection involves a huge number of criteria, so the embedded designers should consider that when they are choosing the best component.
- ✓ Multiple criteria are usually taking place; some of them are quantitative while the others are qualitative.
- ✓ The criteria itself could be conflicting to each other, such as quality against price.
- ✓ Changing in criteria may happen across time and place.
- ✓ Besides the huge number of alternatives may be involved according to the competitiveness among them.

Component selection is a multi-criteria problem which includes both qualitative and quantitative factors. Thus, attention should be given to component selection problem by embedded designers in order to make the right decisions. There are a variety of steps that often embedded designers follow in order to make the right decisions and finally be capable of selecting the most appropriate component. It is agreed that component selection decision is so complicated and difficult to cope with and thus authors proposed a mathematical model in component selection which will help the designers to identify the right components for the new or existing designs.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

II. Related Work

A. Linear Weightage Model

One of the linear weightage models is maximax. This model is very easy and mostly depending upon decision maker's judgment as they have to assign weights to the criteria that involve in decision making process. In most cases there are some criteria considered as more important than others, such as Operating voltage, ADC resolution, ADC Channel number and communication peripheral. Decision makers should assigned weight to each individual criterion in order to determine the relative importance of each one. These weights play a vital role in decision making process and extremely affect the final decision. After identifying all the criteria related to website selection decision, decision maker has to determine threshold for each criterion. In fact, threshold can be divided into two types, i.e. maximum and minimum. One criterion may be "Smaller is better" and the threshold for this type of criteria must be maximum. On the other hand other criteria can be considered as "larger is better" where thresholds must be minimum.

$$Cmax = Max - Component / Max - Min \quad (1)$$

Where,

Cmax = Component value that has maximum type of threshold with respect to a particular attribute/criterion.

Component = Specific component that is considered at the time.

Max = Maximum value of particular attribute/criteria among all component.

Min = Minimum value of the same attribute among the whole component.

In the other case when the attribute is classified under the minimum type of threshold, formula 2 is the only option for calculating the component's value.

$$Cmin = Component - Min / Max - Min \quad (2)$$

Where.

Cmin = Component value that has minimum type of threshold with respect to a particular attribute/criterion.

Component = Specific component that is considered at the time.
Max = Maximum value of particular attribute/criteria among all component

Min = Minimum value of the same attribute among the whole component.

The idea of using formula 1 and formula 2 is extremely valuable because they provide a method that enables the comparisons among decision criteria. Usually decision criteria have different units of measure so any comparisons among those criteria are not logically acceptable. By using the data normalization concept which was represented in formula 1 and formula 2, all the criteria will be having weights instead of a variety of measurement units and then the comparisons can simply be made. When all values of the criteria matrix are calculated, series of calculations should be achieved by multiplying weights Wi of criteria by the whole values Xi within the matrix. The total score should also be calculated using formula 3 for each component which represents the components scores. The final decision table includes a total score for each component and the one who gains the highest score is recommended as the best component over all. The limitation of this model is assigning weights to various criteria.

$$\text{Total Score} = \Sigma W_i X_i \quad (3)$$

B. Analytic Hierarchy Process

The Analytical Hierarchy Process Model was designed by TL Saaty [3] as a decision making aid. The Analytic Hierarchy Process is based on the assumption that when faced with a complex decision the natural human reaction is to cluster the decision elements according to their common characteristics.

In AHP the problems are usually presented in a hierarchical structure and the decision maker is guided throughout a subsequent series of pairwise comparisons to express the relative strength of the elements in the hierarchy. In general the hierarchy structure encompasses of three levels, where the top level represents the goal, and the lowest level has the component under consideration. The intermediate level contains the criteria under which each component is evaluated.

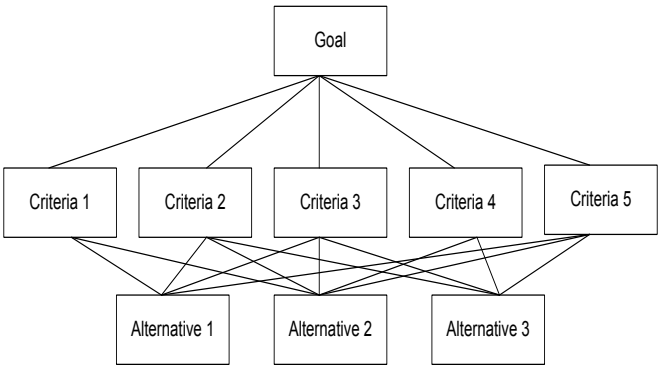


Fig. 1. Analytical Hierarchy Process Model

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

III. Proposed Hybrid Model

Based on the previous discussion about both models, there is an urgent need for new model that can support the component selection decision and offer a powerful tool which can ultimately produce satisfactory results. This paper intends to achieve this objective by proposing a new hybrid model. This new model concentrates on avoiding all the shortcomings mentioned above. It combines two different aspects from both AHP and linear weightage model.

The new model uses the measurement scale of AHP model to determine to which degree each single criterion is preferred in comparison with others. Once the pairwise comparisons have been made, decision maker can obtain the weights of the whole criteria when the relative preference of criteria is specified. The next step in the proposed model is to assign thresholds to all criteria considering "larger is better" or "smaller is better".

First stage is to obtain preference criteria matrix, by means of identifying various criteria against each other. Make pairwise comparison between the criteria by assigning weights in 1-9 scale. By performing three steps like sum the elements in each column, divide each value by its column total and calculate row averages. Finally by doing all the three steps we can obtain weigtages of each criterion. The second stage is to apply linear weightage model by finding the thresholds from the original component data and after normalization process by multiplying the weights obtain from the above process, we can get the final decision table matrix. Calculation of the whole values in the decision table matrix has to be produced by considering the two formulae. If the threshold is maximum then formula 1 should be used, otherwise formula 2 is applied for minimum threshold. When the whole cells that represent each component across only criteria will be filled with a certain value in the decision table matrix, then each column will multiply by the column of criteria weights and obtain the new values of these cells. Now each column represents one of the competitive components, the last step in the proposed model is to compute the sum of each column to get the final scores of all components. The highest score indicates to the best component and that component will be recommended as the most appropriate component among the competitive components.

IV. Numerical Illustration

The data for this case study have been collected from the microcontroller selection study for the project Design and Development of Object Tracking system for environmental sensitive object in transit.

First row in Table I shows the selection criteria for the microcontroller. These criteria which are involved in the component selection process are eight different criteria which describe each product. The columns represent the twelve competitive products.

TABLE I. Micrcontroller Technical Specifications

#	Microcontroller	CPU	Power consumption	Flash	EEPROM	RAM	Min Operating Voltage	USB	RTC	Expertize Level	Pins
Units		Bit	µW	Kb	Bytes	Bytes	Volts	Yes/No	Yes/No	High/Low	No.
1	PIC18LF14K50	8	10.8	16	256	768	1.8	Yes	No	High	20
2	PIC16LF1829	8	12.6	8	256	1024	1.8	No	No	High	20
3	PIC18F87K90	8	9.9	128	1024	4096	1.8	No	Yes	High	80
4	PIC24FJ32GB004	16	30	64	0	8192	2.0	Yes	Yes	High	44
5	PIC18LF26J50	8	12.4	64	0	3776	2.0	Yes	Yes	High	24
6	MSP430F2013	16	17.28	2	256	128	1.8	No	No	Low	14
7	MSP430F5528	16	11.7	128	0	8192	1.8	Yes	Yes	Low	80
8	STM8L152M8	8	56	64	2048	4096	1.65	No	Yes	Low	80
9	STM32L15xVx	32	45	128	4096	16384	1.8	Yes	Yes	Low	48
10	MC9S08JE128	8	126	128	0	12288	1.8	Yes	No	Low	64
11	MC9S08MM128	8	126	128	0	12288	1.8	Yes	No	Low	64
12	PIC24F16KA102	16	14.4	16	512	1536	1.8	No	Yes	High	20

The ten criteria for the selection of microcontroller are CPU architecture, Typical Power consumption at 32 KHz with VDD = 1.8 v, Flash, EEPROM, RAM, Minimum operating voltage, USB support, availability of RTC, Expertise level and number of pins. Table II is prepared using the formula number 1 and 2 and is named as base reference values.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

TABLE II. NORMALIZE COMPONENT VALUES MATRIX

#	Microcontroller	Min	Max	Min	Min	Min	Min	Min	Min	Min	Max
1	PIC18LF14K50	1.00	0.99	0.11	0.06	0.04	0.57	1.00	0.00	1.00	0.91
2	PIC16LF1829	1.00	0.98	0.05	0.06	0.06	0.57	0.00	0.00	1.00	0.91
3	PIC18F87K90	1.00	1.00	1.00	0.25	0.24	0.57	0.00	1.00	1.00	0.00
4	PIC24FJ32GB004	0.67	0.83	0.49	0.00	0.50	0.00	1.00	1.00	1.00	0.55
5	PIC18LF26J50	1.00	0.98	0.49	0.00	0.22	0.00	1.00	1.00	1.00	0.85
6	MSP430F2013	0.67	0.94	0.00	0.06	0.00	0.57	0.00	0.00	0.00	1.00
7	MSP430F5528	0.67	0.98	1.00	0.00	0.50	0.57	1.00	1.00	0.00	0.00
8	STM8L152M8	1.00	0.60	0.49	0.50	0.24	1.00	0.00	1.00	0.00	0.00
9	STM32L15xVx	0.00	0.70	1.00	1.00	1.00	0.57	1.00	1.00	0.00	0.48
10	MC9S08JE128	1.00	0.00	1.00	0.00	0.75	0.57	1.00	0.00	0.00	0.24
11	MC9S08MM128	1.00	0.00	1.00	0.00	0.75	0.57	1.00	0.00	0.00	0.24
12	PIC24F16KA102	0.67	0.96	0.11	0.13	0.09	0.57	0.00	1.00	1.00	0.91

The Pairwise comparison preference Criteria Matrix is prepared using the Analytic Hierarchy Process. CPU, Flash, EEPROM and RAM have an equal preference of criteria that is why each of them is filled with ones. However as other criteria's has high priority appropriately cells are filled with 1/3, 1/5 and 1/7.

TABLE III. PAIRWISE COMPARISON PREFERENCE CRITERIA MATRIX

	CPU	Power Consumption	Flash	EEPROM	RAM	Minimum Operating Voltage	USB	RTC	Expertise Level	Pins
CPU	1	1/7	1	1	1	1	1/3	1/3	1/5	1/3
Power Consumption	7	1	7	7	7	7	5	5	3	5
Flash	1	1/7	1	1	1	1	1/3	1/3	1/5	1/3
EEPROM	1	1/7	1	1	1	1	1/3	1/3	1/5	1/3
RAM	1	1/7	1	1	1	1	1/3	1/3	1/3	1/3
Minimum Operating Voltage	1	1/7	1	1	1	1	1/3	1/3	1/5	1/3
USB	3	1/5	3	3	3	3	1	1/3	1/5	1/3
RTC	3	1/5	3	3	3	3	1	1	3	1
Expertise Level	5	1/3	5	5	5	5	1	1	1	1
Number of Pins	3	1/5	3	3	3	3	1	1	1	1
Total	26.00	2.65	26.00	26.00	26.00	26.00	10.67	10.00	9.33	10.00

The next step is to obtain the weight for each criterion by normalized the data in Table III. The process follows three major steps, which are as below

a) Sum the elements in each column.
b) Divide each value by its column total.
c) Calculate row averages.

Performing the above steps on the data mentioned in Table III yields the normalized matrix of criteria as illustrated in Table IV. The average weights of rows are computed in the last column to indicate the weights of the criteria.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

TABLE IV. WEIGHTS OF EACH COMPONENT

	CPU	Power Consumption	Flash	EEPROM	RAM	Minimum Operating Voltage	USB	RTC	Expertise Level	Pins	Weight
CPU	0.0385	0.0540	0.0385	0.0385	0.0385	0.0385	0.0313	0.0333	0.0214	0.0333	0.0366
Power Consumption	0.2692	0.3777	0.2692	0.2692	0.2692	0.2692	0.4688	0.5000	0.3214	0.5000	0.3514
Flash	0.0385	0.0540	0.0385	0.0385	0.0385	0.0385	0.0313	0.0333	0.0214	0.0333	0.0366
EEPROM	0.0385	0.0540	0.0385	0.0385	0.0385	0.0385	0.0313	0.0333	0.0214	0.0333	0.0366
RAM	0.0385	0.0540	0.0385	0.0385	0.0385	0.0385	0.0313	0.0333	0.0357	0.0333	0.0380
Minimum Operating Voltage	0.0385	0.0540	0.0385	0.0385	0.0385	0.0385	0.0313	0.0333	0.0214	0.0333	0.0366
USB	0.1154	0.0755	0.1154	0.1154	0.1154	0.1154	0.0938	0.0333	0.0214	0.0333	0.0834
RTC	0.1154	0.0755	0.1154	0.1154	0.1154	0.1154	0.0938	0.1000	0.3214	0.1000	0.1268
Expertise Level	0.1923	0.1259	0.1923	0.1923	0.1923	0.1923	0.0938	0.1000	0.1071	0.1000	0.1488
Number of Pins	0.1154	0.0755	0.1154	0.1154	0.1154	0.1154	0.0938	0.1000	0.1071	0.1000	0.1053
Total	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

TABLE V. WEIGHT AND COMPONENT VALUES MATRIX

#	Microcontroller	CPU	Power consumption	Flash	EEPROM	RAM	Min Operating Voltage	USB	RTC	Expertize Level	Pins	Score
Weight		**0.0366**	**0.3514**	**0.0366**	**0.0366**	**0.0380**	**0.0366**	**0.0834**	**0.1268**	**0.1488**	**0.1053**	
1	PIC18LF14K50	0.0366	0.3487	0.0041	0.0023	0.0015	0.0209	0.0834	0.0000	0.1488	0.0958	**0.74**
2	PIC16LF1829	0.0366	0.3432	0.0017	0.0023	0.0021	0.0209	0.0000	0.0000	0.1488	0.0958	**0.65**
3	PIC18F87K90	0.0366	0.3514	0.0366	0.0091	0.0093	0.0209	0.0000	0.1268	0.1488	0.0000	**0.74**
4	PIC24FJ32GB004	0.0244	0.2906	0.0180	0.0000	0.0188	0.0000	0.0834	0.1268	0.1488	0.0575	**0.77**
5	PIC18LF26J50	0.0366	0.3438	0.0180	0.0000	0.0085	0.0000	0.0834	0.1268	0.1488	0.0894	**0.86**
6	MSP430F2013	0.0244	0.3291	0.0000	0.0023	0.0000	0.0209	0.0000	0.0000	0.0000	0.1053	**0.48**
7	MSP430F5528	0.0244	0.3460	0.0366	0.0000	0.0188	0.0209	0.0834	0.1268	0.0000	0.0000	**0.66**
8	STM8L152M8	0.0366	0.2119	0.0180	0.0183	0.0093	0.0366	0.0000	0.1268	0.0000	0.0000	**0.46**
9	STM32L15xVx	0.0000	0.2452	0.0366	0.0366	0.0380	0.0209	0.0834	0.1268	0.0000	0.0511	**0.64**
10	MC9S08JE128	0.0366	0.0000	0.0366	0.0000	0.0284	0.0209	0.0834	0.0000	0.0000	0.0255	**0.23**
11	MC9S08MM128	0.0366	0.0000	0.0366	0.0000	0.0284	0.0209	0.0834	0.0000	0.0000	0.0255	**0.23**
12	PIC24F16KA102	0.0244	0.3378	0.0041	0.0046	0.0033	0.0209	0.0000	0.1268	0.1488	0.0958	**0.77**

V. CONCLUSION

The proposed hybrid model is considered as a robust tool that can assist decision maker in the process of component selection. In addition, the proposed model saves time because there are only a few computations to be done. This model is easy to understand and easy to use. Also it saves effort due to its simplicity, and that will strongly accelerate the component selection decision as well as improve the whole business processes within organizations in turn.

Other advantage of the proposed model is avoiding the limitation in the linear weightage model which assigns the weights of criteria directly by decision maker based on their experience and gut feeling. The proposed model uses the AHP pairwise comparisons and the measurement 1-9 scale to generate the weights for the criteria. This method provides good solution when compared to human judgment. Thus the proposed model overcomes the absolute dependency on human judgment as in the case of Linear Weightage model.

In conclusion, the proposed model can be considered as a powerful model for component selection problem. It fully

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

integrates the advantages of both linear weightage model and AHP approach.

ACKNOWLEDGMENT

This work was done as a part of project titled "Design and Development of Object Tracking system for environmental sensitive object in transit" funded by Department of Information Technology (DIT) Ministry of Communications and Information Technology, Government of India. Authors are thankful to Dr. Debashish Dutta (GC – R & D in IT Group) and Smt. Geeta Kathpaliya (Director) for the support. The authors are indebted to Dr. George Varkey, Executive Director C-DAC Noida to give enough space and freedom to cultivate and nurture the research areas in embedded systems.

REFERENCES

[1] Michael G. Pecht, 2004, Parts Selection and Management : John Wiley & Sons, Inc
[2] General Specification for Microcircuits, Rev. J, MIL-M-38 510, 1991.
[3] Saaty T. L, 1980, The analytic hierarchy process: planning, priority setting, resources allocation. London: McGraw-Hill.
[4] Tianbiao Yu, Jing Zhou, Kai Zhao,. "Study on Project Experts' Evaluation Based on Analytic Hierarchy Process and Fuzzy Comprehensive Evaluation ", International Conference on Intelligent Computation Technology and Automation (ICICTA), vol. I, pp.941-945, 2008
[5] Wei-kang Wang, Wu Wen, W. B Chang, and Hao- Chen Huang, "A knowledge-based decision support system for government vendor selection and bidding", JCIS-2006 proceeding, 2006.
[6] Dongjoo lee, Tachee lee, sue-kyung, ok-ran jeong, Hyenosang EOM, and Sang-goo lee," Best choice: a Decision Support System for Supplier Selection in e- Marketplace", Verlage Berlin Heidelberg, 2006.
[7] E. Gonza´lez and G. Quesada, "Determining the importance of the supplier selection process in manufacturing: a case study", International Journal of Physical Distribution & Logistics Management, Vol. 34, No. 6, 2004.
[8] Dan Wang, Yezhuang Tian, and Yunaquan Hu, "Empirical study of supplier strategies across the study chain management in manufacturing companies", IEEE, vol.1, 2004, pp 85-89.
[9] B. S. Sahay, and A. K. Gupta, "Development of software selection criteria for supply chain solutions", *Industrial Management and Data Systems*, vol. 103, no. 2, 2003; pp. 97-110.
[10] W. Wen, W. K. Wang, and T. H. Wang, "A hybrid knowledge-based decision support system for enterprise mergers and acquisitions", *Expert Systems with Applications*, vol. 28, no. 3, 2005, 569-582.
[11] Marvin E. G, Gioconda Quesada, and Carlo, 2004, "Determining the importance of supplier selection process in manufacturing: A case study", International journal of physical distribution & logistic management, Vol.34, No.6, pp.492-504.
[12] Russell, Roberta S. and Taylor III, Bernard W. Operations Management 4th edition. Upper Saddle river, New Jersey: Prentice Hall, 2003.

AUTHORS PROFILE

Ashutosh Gupta holds Bachelors in Electronics & Communication from Visveswaraiah Technological University, Belgaum, India and Post-Graduation in Telecommunication Network Planning and Management from Indian Institute of Technology, Kharagpur (IIT – Kgp). As a part of work integrated program he has completed M.S. (Masters of Science) in Quality Management from BITS Pilani. Presently he is working as Technical Officer in Embedded Systems group at C-DAC, before joining the present assignment he was with Wipro Technologies as Senior Project Engineer. His interest covers the areas of RFID, Sensor networks and HVAC systems. He has several national and International publication and Patent in Embedded domain to his credit.

Chandan Maity received his Bachelors of Engineering in Electrical Engineering from Burdwan University, West Bengal, India. Presently he is working as Senior Technical Officer in Embedded Systems group at C-DAC. From 2004 to Aug, 2006 he was with Wartsila India Limited as Electrical Engineer. From Aug, 2006 to Dec, 2006 he was with IIT Kanpur as Research Associate. From Dec, 2006 to Nov, 2007 he was the R&D and Technical Head in Iaito Infotech Pvt. Ltd. His interests cover the domain of RFID, GSM, AI, Ubiquitous system. He has several national and International publication and Patent in Embedded domain to his credit.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

Detection and Elimination of Ocular Artifacts from EEG Data Using Wavelet Decomposition Technique

Shah Aqueel Ahmed, D .Elizabath Rani, Syed Abdul Sattar

***Abstract*--This paper presents detection and elimination of ocular artifact from electroencephalographic data using stationary wavelet transform. Usually all the biomedical signals are contaminated with the noise. This noise source increases the difficulty in analyzing the EEG signal. In this paper we are dealing with the EEG signal contaminated with ocular artifacts. Ocular artifacts are more predominant over other artifacts. Since, these ocular artifacts occupy lower frequencies they are difficult to eliminate. Stationary wavelet transform and its inverse are applied in this paper for detection and elimination of ocular artifact.**

***Index Terms*--EEG (Electroencephalography), OA (ocular artifact), SWT (Stationary Wavelet Transform) and EOG (Electrooculography).**

I. Introduction

Electroencephalogram is a valuable tool for clinicians in numerous applications, from the diagnosis of neurological disorders, to the clinical monitoring of depth of anesthesia. Eye movement and blink produce electrical signals around the eye which spread across the scalp and contaminates the EEG. These contaminating potentials are commonly referred to as ocular artifacts (OA's) [1].

At present there are three main methods for artifact processing and they are

1. Artifact rejection(elimination of an artifact contaminated section of EEG)
2. Artifact minimization (nulling, canceling or subtracting of artifacts)
3. Artifact clustering(grouping of artifacts as a particular type of "EEG activity")

In artifact rejection method, the epochs contaminated with artifacts (OA) are rejected this leads to substantial loss of valuable data, because of which EEG cannot be completely monitored and hence cannot diagnose the diseases properly [2].

Shah Aqueel Ahmed, and Dr. Syed Abdul Sattar are with Royal Institute of Technology & Science, Hyderabad – 501503, India (email: shah_aqueel@rediffmail.com).

Dr. D. Elizabath Rani is with Gitam Institute of Engineering & Technology, GITAM University, Vishakapatnam, AP, India.

Artifact clustering is the special case of the artifact rejection, with the advantage that specific methods for rejection of each type of artifact are not required Artifact minimization techniques are preferable in general to artifact rejection techniques for the same artifact, since no loss of data is entailed. Various other methods have been proposed for correcting ocular artifacts and are discussed in brief. Other attempts have been made on different methods based on regression in time domain or frequency domain techniques for removing OA's. Regression methods whether in time or frequency domain depend on having one or more regression (EOG) channel. Also both these methods share an inherent weakness that spread of excitation from eye movements and EEG signal is bidirectional. Therefore regression based artifact removal eliminates the neural potentials common to reference electrodes and to other frontal electrodes [3].

Another class of methods is based on a linear decomposition of the EEG and EOG leads to source components identifying artifactual components and then reconstructing the EEG without the artifactual components. Principal component analysis (PCA) was introduced to remove the artifacts from the EEG. It outperformed the regression based method. However, PCA cannot completely separate OA from EEG, when both the waveforms have similar voltage magnitudes.PCA decomposes the lead into uncorrelated, but not necessarily independent components that are spatially orthogonal and thus it cannot deal higher order statistical dependencies. An alternate approach is to use independent component analysis(ICA),which was developed in the context of blind source separation problems to obtain components that are approximately independent.ICA has been used to correct for ocular artifacts ,as well as artifacts generated by other sources. ICA is an extension of PCA which not only decorrelates but can also deal with higher order statistical dependencies. ICA algorithms are superior to PCA in removing a wide variety of artifacts from the EEG even in the case of comparable amplitudes [4].

II. Wavelet Decomposition Technique

Mathematical transformations are applied to the signals to obtain the further information from that signal that is not readily available in the raw signal. In this paper we assume that a time domain signal, as a raw signal and a signal that has been transformed by any of the available mathematical

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

transformation, as a processed signal. Most of the signals in practice are time domain signals in their raw format. This representation is not always the best representation of the signal, for most signal processing applications. In many cases, the most distinguished information is hidden in the frequency content of the signal. There are number of transformations that can be applied among which the Fourier transforms are probably by far the most popular but Fourier analysis has a serious drawback in transforming to the frequency domain, time information is lost. To overcome this, short time fourier transform was introduced .The short time fourier transform (STFT) represents a sort of compromise between the time and frequency based views of a signal. It provides the information about both when and at what frequencies a signal event occurs. However, this information can be obtained with limited precession and that precession is determined by the size of the window. Wavelet analysis represents the next logical step: A windowing technique with variable sized regions, Wavelet analysis allows the use of long time intervals where we want more precise low frequency information and shorter regions where we want high frequency information [5].

In this paper we are concerned with EEG signal, since the EEG signal is not a stationary signal and it is also an unpredicted signal, therefore we are going with discrete wavelet transform. In this method we are decomposing the EEG signal up to 8 levels using symlet 3 filters.

III. The Process Of Selecting The Threshold

Ocular artifacts are large, transient, slow waves. They occupy lower frequency range i.e, from 0Hz to 6-7Hz for the eye movement artifacts and typically up to the alpha band (8-13Hz), excluding very low frequencies, for the eye blink. When compared with the uncontaminated EEG,the amplitudes of the OA'sare of much higher order.

In the awake conscious state neurons are firing in a more independent fashion, as a result of this desyncronization, the awake EEG signal is even more random spacing. The true EEG is a noise like signal. Therefore any clear patterns cannot be observed within it, nor can we simply correlate the particular underlying events with its shape. Therefore the EOG can be removed by recovering the regression function from the recorded EEG.A wavelet decomposition technique is a simple and an effective technique for denoising.[7]

The EEG recorded is the combination of true EEG signal and the external noise. This external noise may be due to different artifacts , ,and this is denoted as k(t).The true EEG can be denoted as E(t).therefore the measured signal can be represented as

$$X(t) = E(t) + K(t) \quad \text{---------- (1)}$$

In this paper we assume that E(t) and K(t) are not correlated. Thresholding is a technique used for denoising both the signal and image. Selecting an appropriate threshold limit is the difficult part in this process. The formula used for this thresholding is as follows.

$$T = 0.25*\max(er) \quad \text{---------- (2)}$$

Where max(er) is the maximum value in the low frequency band. The EEG signal is decomposed using wavelet decomposition technique up to 8 levels. After decomposing the signal up to 8 levels we are left with approximate and detailed coefficients. Approximate coefficients are the low frequency component which has to be discarded; where as detailed coefficients are high frequency components which are to be restored, after comparing them with the calculated threshold. As we have discussed previously OA's occupy lower frequencies so we are only concerned with low frequency components. The choice of threshold limit should be such that it should not remove the original signal coefficients leading to the loss of EEG data.

IV. Methodology

In this paper we are presenting a technique based on wavelet decomposition for the removal of the ocular artifacts. For this purpose we have taken EEG data of 8 channels. First of all we are decomposing the data of the first channel upto 8 levels using symlet 3 filter, next we are calculating the threshold, then comparing each coefficients with the threshold and keeping only those coefficients larger than threshold and applying wavelet reconstruction to obtain the estimated EEG signal. This process is repeated for all the remaining channels [11].

V. Results

Figures of all the 8 channels are given one by one by plotting both the contaminated and corrected EEG.As we have mentioned that the amplitude of ocular artifact will be much larger than the original EEG signal which is clearly seen in the graphs of all the 8 channels.

Channel 1:

In the contaminated EEG signal of first channel we can observe a peak in between 50th and 100th sample. This peak is identified as ocular artifact in EEG signal. As we can observe that the amplitude of the Peak is above 200µv, and the amplitude of corrected EEG is reduced to a little above 50µv.

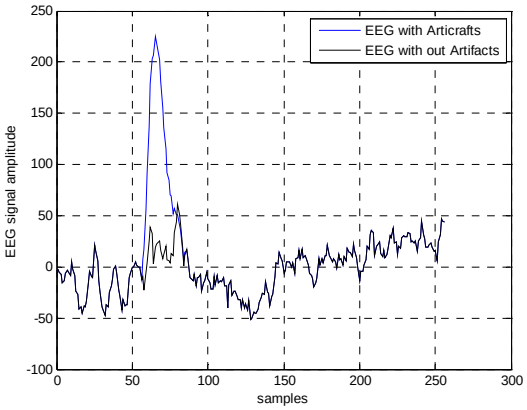


Fig.1. Combination of contaminated and corrected EEG of channel1

Channel 2:

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

In the contaminated EEG signal of second channel we can observe a peak in between 50th and 100th sample. This peak is identified as ocular artifact in EEG signal ,As we can observe that the amplitude of the Peak is About 96µv.After applying wavelet decomposition technique the amplitude of EEG signal is reduced to a about 35µv,which is called corrected EEG signal.

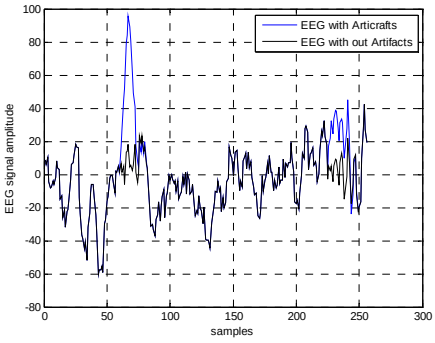


Fig. 2. Combination of contaminated and corrected EEG signal of channel2

Channel 3:

In the contaminated EEG signal we can observe a peak in between 50th and 100th sample. This peak is identified as ocular artifact in EEG signal which is recorded in the third channel .the amplitude of the Peak is above 80µv, the amplitude of corrected EEG is reduced to a about 20µv.

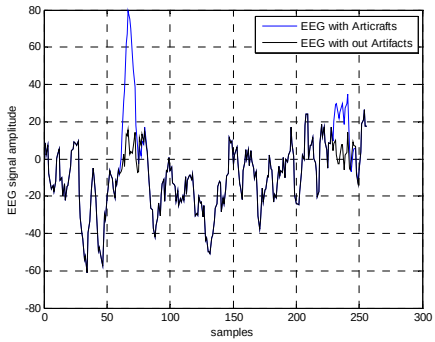


Fig. 3. Combination of contaminated and corrected EEG signal of channel 3

Channel 4:

In the contaminated EEG signal we can observe a peak in between 50th and 100th sample. This peak is identified as ocular artifact in EEG signal which is recorded in the fourth channel. As we can observe that the amplitude of the Peak is about 117µv and after correcting it has reduced to a little above 20µv.

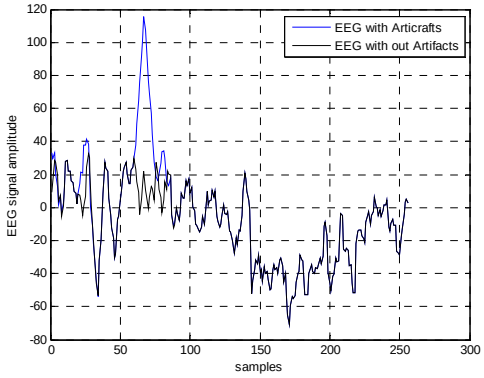


Fig. 4. Combination of contaminated and corrected EEG signal of channel 4

Channel 5:

In contaminated EEG signal we can observe a peak in between 50th and 100th sample. This peak is identified as ocular artifact in EEG signal which is recorded in the fifth channel. As we can observe that the amplitude of the Peak is about 80µv and after correcting it has reduced to 20 µv.

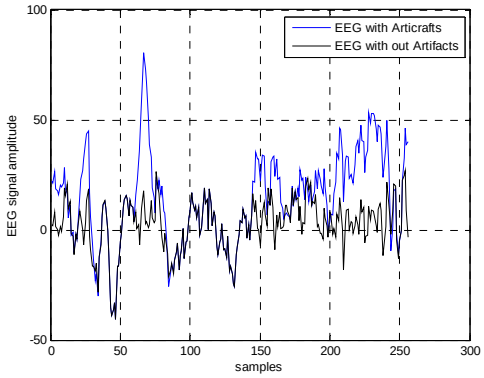


Fig. 5. Combination of contaminated and corrected EEG signal of channel 5

Channel 6:

In contaminated EEG signal we can observe a peak in between 50th and 100th sample. This peak is identified as ocular artifact in EEG signal which is recorded in the sixth channel. As we can observe that the amplitude of the Peak is at 80µv and after correcting it has reduced to 20 µv.

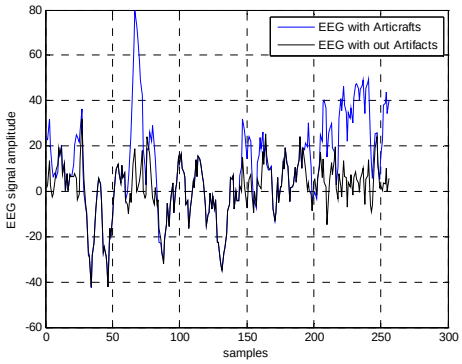


Fig. 6. Combination of contaminated and corrected EEG signal of channel 6

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

Channel 7:

In contaminated EEG signal we can observe a peak in between 50th and 100th sample. This peak is identified as ocular artifact in EEG signal which is recorded in the seventh channel. As we can observe that the amplitude of the Peak is little above 100µv and after correcting it has reduced to about 20 µv.

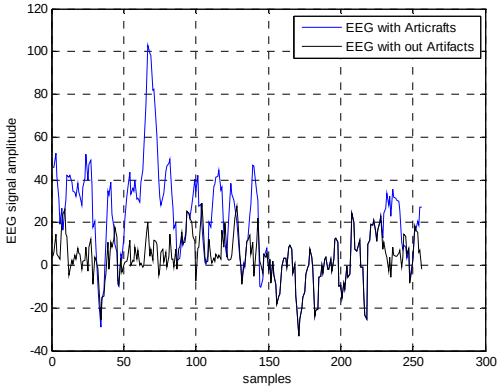


Fig. 7. Combination of contaminated and corrected EEG signal of channel 7

Channel 8:

In contaminated EEG signal we can observe a peak in between 50th and 100th sample. This peak is identified as ocular artifact in EEG signal which is recorded in the eighth channel. As we can observe that the amplitude of the Peak is about 75µv and after correcting it has reduced to18 µv

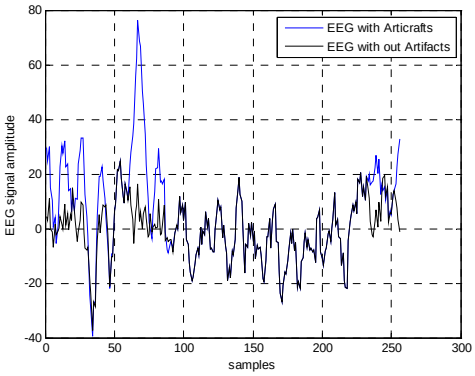


Fig. 8. Combination of contaminated and corrected EEG signal of channel 8

VI. References

[1] Prof.Shah Aqueeel Ahmed.” Studies in EEG for epilepsy, different activities and artifacts.

[2] Prof.Shah Aqueel Ahmed,Prof Mateenuddin H.Quazi,Dr.Syed Abdul sattar ”Detection and elimination of artifacts in Electroencephalographic data”. International Conference on Systemics,Cybernitics and Information.2004

[3] Tatjana Zikov,Stephane Bibian,Guy A.Dumont,Mihai Huzmezan,Craig R.Ries,”A wavelet based denoising technique for ocular artifact correction of the encephalogram”proceedings of the second joint EMBS/BMES conference,2002

[4] P.Senthil kumar, R.Arumughanathan, K.Sivakumar,C.Vimal,”A wavelet based statistical method for denoising of ocular artifacts in EEG signals,IJCSNS International journal of computer science and network security,VOL8 No9,september 2008.

[5] S.Salivahana, A.Vallavaraj & C.Gnanapriya, ”Digital Signal Processing”.

[6] Wills J.Tompkins,”Biomedical Digital Signal Processing”.

[7] Prof.S.G.Kahalekar,Sampat. P, A.G.Shah”DSP applications in biomedical engineering”,ISTE Sponsered Summer School on “Digital signal processing”at SGGSC&T,Nandeed.

[8] R.S.Khandpur, “Biomedical instrumentation”. Second edition, 2003

[9] Dr.M..Arumugum”Biomedical Instrumentation”.

[10] Joseph J .Carr & John M.Brown,”Introduction to Biomedical Equipment technology

[11] Robi Polikar” The wavelet tutorial “ Ames.Iowa 1996

[12] The mathworks Inc, M.A.,”MATLAB user’s guide”. 1997

[13] Rudra Pratap” Getting started with MATLAB 7” 2006.

[14] Webster J.G.,”Medical Instrumentation”.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

Cluster-Based Routing Protocol To Improve Qos In Mobile Adhoc Networks

Prof. M.N. Doja
Department of Computer Engineering
Faculty of Engineering & Technology
Jamia Millia Islamia, New Delhi, India

Mohd. Amjad
Department of Computer Engineering
Faculty of Engineering & Technology
Jamia Millia Islamia, New Delhi, India

Abstract: An Ad Hoc network is a collection of wireless mobile hosts dynamically forming a temporary network without the aid of any existing established infrastructure. Quality of Service (QoS) is a set of service requirements that needs to be met by the network while transporting a packet stream from a source to its destination. QoS support for Mobile Adhoc Networks (MANETs) is a challenging task due to the dynamic topology and limited resources. Characteristics of Mobile Ad Hoc Networks (MANETs) such as lack of central coordination, mobility of hosts, and limited availability of resources make Quality of Service (QoS) provisioning very challenging. Limited resource availability such as battery power, average energy consumption of the network by all of the nodes and insecure medium are some of the major QoS issues to be dealt with. In this paper we have suggested a clustering of participating nodes with minimum energy consumption by the overall network by hierarchical cluster-based routing Algorithm. In this algorithm we have introduced a new metric, next hop availability, which is a combination of two metrics. It maximizes path availability and minimizes travel time of packets and therefore offers a good balance between selection of fast paths and a better use of network resources with minimum energy consumption. In the conclusion it provides simulation result to evaluate the performance on a network simulator.

Keywords : - Power saving protocol, clusters, Quality of service support, Ad hoc network.

1 INTRODUCTION

In an ad hoc network the mobile nodes agree to serve as both routers and hosts. The nodes can dynamically join and leave the network, frequently without warning, and possibly disrupting communication amongst other nodes. Moreover, the limitations on power consumption imposed by portable wireless radios result in a node transmission range that is typically small relative to the span of the network. This limits the propagation range of a mobile node[7]. In such an environment, it may be necessary for one mobile host to enlist the aid of others in forwarding a packet to its destination. These networks can be formed on the fly, without requiring any fixed infrastructure. As these are infrastructure less networks, each node should act also as a router. These characteristics of MANETs such as lack of central coordination, mobility of hosts, limited availability of resources and insecure medium make Quality of Service (QoS) provisioning very challenging. QoS is usually defined as a set of service requirements that need to be met by the network while transporting a packet stream from a source to its destination(s)[11]. This can be achieved by incorporating quality of service (QoS) metrics such as energy consumption by the network, battery life of the nodes and security measures into the routing decisions as opposed to choosing a shortest path. Efficient resource management mechanisms are required for optimal utilization of this scarce resource i.e. battery power. In ad hoc network this operation is called clustering, giving the network a hierarchical organization. A cluster is a connected graph including a cluster head responsible of the management of the cluster, and (possibly) some ordinary nodes. Each node belongs to only one cluster. Some MANETs, such as mobile military networks or future commercial networks may be relatively large (e.g. hundreds or possibly thousands of nodes). A way to support the increasing number of nodes in MANET is to subdivide the whole network into groups, and then create a virtual backbone between delegate nodes in each group[16]. In ad-hoc network this operation is called clustering, giving the network a hierarchical organization.

2 CLUSTERING IN MANETS

A way to support the increasing number of nodes in MANET is to subdivide the whole network into groups, and then create a virtual backbone between delegate nodes in each group. In ad-hoc network this operation is called clustering, giving the network a hierarchical organization. Several cluster based adaptations has been proposed for existed routing protocols and other protocol as ZRP (zone routing Protocol), CBRP (cluster based protocol) have originally exploited this concept[6][10]. Clustering for security can simplify the management of Certificate Authority in a Public Key Infrastructure (PKI) by affecting the full or a subset of Certificate Authority services to cluster-heads, ensuring in this way the availability of the Certificate Authority.
The Hierarchal organization consists of:
Cluster Head: A cluster head, as defined in the literature, serves as a local coordinator for its cluster, performing inter-cluster routing, data forwarding and so on. In our self-

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

.organized clustering scheme the cluster head only serves the purpose of providing a unique ID for the cluster, limiting the cluster boundaries [3].

Cluster Gateway: A cluster gateway is a non cluster-head node with inter-cluster links, so it can access neighboring clusters and forward information between clusters.
Cluster Member: A cluster member is a node that is neither a cluster head nor a cluster gateway.

3 LIMITATIONS OF EXISTING LGORITHMS

None of the two well known cluster based routing protocols i.e ZRP or CBRP leads to an optimal election of cluster-heads since each deals with only a subset of parameters which can possibly impose constraints on the system. However, a cluster-head may not be able handle a large number of nodes due to resource limitations even if these nodes are its immediate neighbors and lie well within its transmission range. In other words, simply covering the area with the minimum number of cluster-heads will put more burdens on the cluster-heads. On the other hand, a large number of cluster-heads will lead to a computationally expensive system [2][[8]. Although this may result in good throughput, the data packets have to go through multiple hops thus implying high latency.

4 OUR ALGORITHM

Assumptions
1. The network is divided into cluster of nodes with a single clusterhead per cluster.
2. No two clusterheads can be one hop neighbors of each other.
3. Overlapping clusters are connected through Gateway nodes.
4. All the ordinary nodes are one-hop from their cluster heads.
5. Each node that requests for an entry permit must keep track of the respective
Weights broadcasted by the neighbor nodes.
6. Battery power is reduced in proportion to the number of packets sent.

The

4.1 DATA DICTIONARY

W_A	Combined weight of each node A.
PC_{WT}	Minimum weight among all W_A.
PC	Possible Clusterhead.
X[]	Neighbor cluster heads in the transmission range of PC.
CHMsg	Cluster head selection message.
Weight[]	Weights of all neighbor cluster heads in the transmission range of PC.
th1	Threshold value 1 (associated with weights of newly selected cluster head).
th2	Threshold value 2 (associated with weights of newly selected cluster head and existing neighbor cluster heads in the transmission range of PC).
n	Total number of existing cluster heads in the whole network.
c	The total number of existing cluster heads in the whole network whose weights are greater than a specified value.
B	Any neighbor node in the transmission range of node A.

TABLE 1: DATA DICTIONARY

4.2 THE DESIGN APPROACH

In this section, we will describe our proposed clustering algorithm called Cluster Base Algorithm (BCA). BCA is a weight based clustering algorithm which uses a weight computed from a set of parameters to elect cluster-heads [13].
The main basic concepts used to derive the needed parameters are given below:

***Max Value*:** Represents the upper bound of the number of nodes that can simultaneously be supported by a cluster-head. Since mobile nodes have limited resources, therefore they can't handle a great number of Nodes. This value is defined according to the remainder of resources of the cluster-head.

Min Value: represents the lower bound of the number of nodes that belong to a given cluster before proceeding to the extension or merging mechanisms. This value is global and the same for the entire network. The *Min Value* may avoid the complexity due to the management of great number of clusters [18][19].

***D hops Clusters*:** As one hop clusters are too small for large ad hoc networks, therefore BCA creates *D* hops clusters where *D* is defined by the underlying protocol or according to the cluster-head state (busy or not). By the way, the diameter of the cluster can be extended in some situations.

Identity (ID): *It* is a unique identifier for each node in the network to avoid any spoofing attacks or perturbation in the election procedure. We propose to use certificate as identity, therefore we suppose the existence of an online or offline Public Key Infrastructure managing the certificate distribution.

***Weight*:** Each node is elected cluster-head according to its weight which is computed from a set of system parameters.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

The node having the greatest weight is elected as cluster-head.

4.3 ELECTION CRITERIA

The following parameters define the criteria on which BCA rely to elect the cluster-head.

***Trust value*:** it measures how much any node in the network is trusted by its neighborhood. It's defined as the average of trust values received from each neighboring node. In order to compute the trust value, we suppose that every mobile node has an intrusion detection mechanism to determine if a node is considered as trust or not by periodically collecting information about the behavior of each neighbor.

Degree: is the number of neighbors of a given node, within a given radius. This parameter is used to choose as cluster-head the node having the maximum neighbors to serve the more number of nodes.

***Battery power*:** This factor is the capability of a node to serve as long as possible. Since cluster-head has extra responsibility and it must communicate as far/long as possible, thus it must be the most powered node.

The Max Value: as defined above, this parameter is used in the election procedure to elect as cluster-head the node which can handle the maximum of nodes.

Link availability: This is the Number of nodes connected at one time from the CH.

***Stability* (Mobility):** This is a useful parameter when electing the cluster-head.

Stability is defined as the difference between two measures of M_D (mean distance) at *t* and *t-1*, it becomes large when the node goes far from its neighbors or whenever its neighbors are going in other direction than the one taken by the considered node. This value is compared with *D* and a node is considered as most stable if it has the less value of stability. $STA = M_{Dt} - M_{D(t-1)}$

In order to elect the most stable node as cluster-head, avoiding frequent roaming, we have computed the *stability* using the following metrics:

- ***The distance*:** The *distance* between two nodes A,B (D_A,D_B), is the number of hops between them, which can be obtained from the packets sent from one to other, or hello message used in routing protocols. The possibility of obtaining the number of hops between two nodes is evident and simple within all existed routing protocols.
- ***The mean distance*:** This is defined as the average of distances between node A and all its neighbors.

***Weight Factors*:** Each of the previous parameters is called partial weight. Each parameter is affected a weight factor defining its degree of importance for the underlying protocol or the network. Since only a subset of these parameters can be used according to the requirements of the network and the underlying protocol, these factors provide more flexibility and large scale of use to our algorithm. For example trust value may take the great value if the underlying protocol is a key management protocol. Factors are given values between 0 and 1, so that the sum of factors is 1.

***Global Weight*:** using all parameters cited above every node in the network computes its global weight. Depending on this weight a given node can be elected as cluster-head or not.

When cluster formation is to be performed by BCA, the nodes can change their position randomly (moves away from each other) due to mobility. The communication among them may become difficult when they place themselves outside the transmission range (tx) of the node from which data has to be transferred. For this reason, transmission power of each node is required for weight calculation. Mobility produces the randomly changed position of each node. But the rates of data transfer capability (Tr) are not same for all the nodes in a cluster formation procedure. It shows the amount of data can be delivered in a certain period of time by a node to all the other nodes in its transmission range. These two parameters have been considered for overall improvement in performance. Using the following formula we calculate the combined weight Wm for each node m, where

We denote W1 , W2 , W3, W4 ,W5 and w6 the partial weights factors corresponding respectively to Trust value, Degree difference, Battery power , Max Value, Stability (Mobility) and Link Status in such a way that the sum of all the factors are 1. The global weight is computed as follows:

$$Wm = \frac{w1*\Delta m + w2*Dm + w3*Vm + w4*Bm}{(W5*tr + w6*tx)}$$

Where Δm : Degree difference

For each node m , Δm is defined as

$\Delta m = |N - \delta|$ for every node m, where,

δ = The total number of ideal neighbors of node

N = Degree (Number of neighbors) of node m= $\sum\{dist(m_i, m_j) < t_{x\,range}\}$

for mj $\in M, m_j \neq m_i$

BCA finds the sum of the distances, Dm, for every node and its neighbors as

Dm : Mean distance

Dm = Sum of distances

$= \sum\sqrt{(X_{mi} - X_{mj})^2 - (Y_{mi} - Y_{mj})^2}$ for (m_i,m_j $\in M$ and

Average relative distance $\overline{D_m} = D_m / N$

Vm: Average relative speed

$$V_{mi,mj,t} = 1/N \sum_{i=1}^{N} |(Vmi, mj, t)|$$

Relative velocity with its neighbor nodes = $V_{(mi,\ m\ j,\ t)} = V_{(mi,\ t)} - V_{(m\ j,\ t)}$

Running average of the velocity

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

$V_m = 1/[\sqrt{(X_t - X_{t-1})^2 + (Y_t - Yt_{-1})^2}]$

where, T = the time for node m motion from the coordinates of (X_{t-1}, Y_{t-1}) to (Xt, Yt)
at time (t-1) and t respectively.
Bm: Battery Power
The Battery Power, Bm, during which a node m acts as a cluster head is obtained.
Tr : Transmission rate
Tx: Transmission power

$$Tx = Tr\ (\frac{\lambda}{4\pi d})^n G_r G_t$$

Gr and Gt are the antenna gain at the transmitter and receiver end.
n : Path loss exponent
For the ideal condition $G_r = G_t = 1$ and it is supposed that the transmission rate is equal to the bandwidth of the channel i.e. there is full utilization of the bandwidth then $T_r = 1$

$$Pt = \frac{(4\pi d)^2}{\lambda^2} = \frac{(4\pi fd)^2}{c^2}$$

λ : Wavelength in meter
c: speed of the light= $3x10^8$ m/s
f: Frequency in Hertz
d: Distance between transmitter and receiver in meter
In our implementation, the weights w1, w2, w3, w4, w5 and w6 are initialized as follows: w1=0.1, w2=0.2, w3=0.5, w4=0.1, w5=0.05 and w6=0.05.
Data transfer rate (tr) = (C1*60)/T1 packets/min where,
T1 =Packet transfer duration in second.
C1 =Number of packets transferred in T1 seconds.
For each node, the range of transmission is tx α (T2 – T1)/2. At time T1 the HELLO message is at co-ordinate (X_{T1}, Y_{T1}) and after (T2 – T1)/2 time the message reached the co-ordinate $(X_{((T2-T1)/2)} - Y_{((T2-T1)/2)})$.
Thus, the transmission range is

$$(tx) = \sqrt{((X_{((T2-T1)/2} - X_{T1})^2 - (Y_{((T2-T1)/2} - Y_{T2})^2)}$$

T: Specified period of time such that,
$T \geq \{Max\ ((T_2 - T_1)\ for\ all\ N\ nodes)\}$
Each node that wishes to join a cluster must keep information about weights of its neighbor cluster heads. To maximize the resource utilization, we can choose to have the minimum number of cluster heads to cover the whole geographical area over which the nodes are distributed. The whole area can be split up into zones. The size of each zone can be determined by the transmission range of the nodes, selected as cluster heads.

5 INITIALIZATION OF CLUSTER NODES

The Init procedure is executed by each node in a no determinist status. A node with this status is a node which isn't attached yet to any cluster, this may be caused by a link failure, a roaming, or whenever a node coming for the first time to the network.
First the node listens if there is any neighboring cluster-head CH (Line 3). If this is the case it chooses the nearest cluster and joins it (Line 11, 12). Otherwise it launches the election procedure to elect a new CH in this neighborhood (Line 6).

```
Procedure Init ()
1. Begin
2. If status=N then
3. CH-list=Get_beacons();
4. If CH-list=null then
5. begin
6. Election ( );
7. exit();
8. end;
9. else
10. begin
11. CH=CH-list.Get_Nearest ( );
12. JOIN(CH);
13. end;
14. exit;
15. End;
```

6 GENERATING TRAFFIC MODEL OF BCA

Random traffic connections of CBR can be setup between mobile nodes using a traffic-scenario generator script. This traffic generator script is available under Glomosim-2.03/Glomosim/bin/BCA.pc and is called BCA.pc. It can be used to create CBR traffics connections between wireless mobile nodes. So the command line looks like the following:

```
Glomosim bin config.in  [-type cbr|BCA] [-nn nodes] [-seed seed] [-mc connections] [-rate rate] > [file name ]
```

For the simulations carried out, traffic models were generated for 20, 30, 40 & 50 nodes with cbr traffic sources, with maximum connections of 20,30, 40 & 50 nodes at a rate of 8kbps.

Mobility Models

The node-movement generator is available under Glomosim-2.03/Glomosim/bin/GlomoMain/jvac
*.java/java_gui directory and consists of setdest {BCA.pc} and Makefile.

Mobility models were created for the simulations using 20 and 40 nodes with pause times of 0, 4,8,16 and 24 seconds, maximum speed of 20m/s, topology boundary of 500x500 and simulation time of 500secs.
Evaluating Packet delivery fraction (pdf) Count the number of sent packets and number of received packets from the file (.stat file). By these two types of packets, we can calculate pdf for each trace file using the formula given below:Packet delivery fraction (pdf %) = (received packets/ sent packets) *100

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

Evaluating Routing Load

Routing load is the ratio of routing packet sent divide by routing packets receives, i.e.,

Routing Load= (routing packets sent) / received packets

7 ENERGY CALCULATION IN BCA

Batteries are the major source of energy in mobile nodes. To provide greater portability, batteries need to be small and lightweight, which unfortunately restricts the total energy that they can carry. Once batteries exhaust their energy, they need to be replaced or recharged, which typically reduces the independence of a mobile node to a few hours of operation. Energy consumption, in communication-related tasks, depends on the communication mode of a node. A node may be transmitting, receiving, or in idle mode. Naturally, transmission consumes more energy than the other two modes. From the routing perspective, our interest is in selecting routes in such a way that the transmission and reception of packets is intelligently distributed on the network so as to maximize the overall average battery lifetime of the nodes. Therefore, we are interested in getting forward cluster_head agents to select, with greater frequency, those nodes which have the longest remaining battery lifetime.

$$B^{r}_{\min}(t) = \min_{i \in N_r} B_i(t)$$

$B_i(t)$ is the residual of battery power of node I at time t, $B^r_{min}(t)$ as the minimum residual energy power of the nodes along route r.

Let $B_{max}(t)$ and $B_{min}(t)$ be the maximum and minimum values among all $B^r_{min}(t)$ in the route then

$$B_{\max}(t) = \max_{r \in R} B^{r}_{\min}(t) \quad \text{and}$$

$$B_{\min}(t) = \min_{r \in R} B^{r}_{\min}(t)$$

For each node r , let E_r denote the energy required by the transmitting nodes the E_{min} is the minimum energy among all E_r i.e.

$$E_{\min} = \min_{r \in R} E_r$$

We use E_r- E_{min} to define how efficiently the route r uses the energy. To save energy this value should be as small as possible.

There is at least one route in the cluster set, whose B^r_{min}= B_{max} will be always in cluster. This means that there is always a route to be chosen from the clusterhead.

If the nodes batteries' remaining energy is not considered in the optimization, the best path's node energy will be used unfairly more than the other nodes in the network. These nodes may fail after a short time because of their battery depletion, whereas other nodes in the network may still have high energy in their batteries.

For the simulator GloMoSim we can write the code in the following path:

"glomosim2.03\glomosim\radio\radio_accnoise.pc"and follow the given instruction:

```
accnoise->stats.energyConsumed
    += txDuration *
(BATTERY_TX_POWER_COEFFICIENT
            * thisRadio->txPower_mW
            + BATTERY_TX_POWER_OFFSET
            - BATTERY_RX_POWER);
accnoise->stats.energyConsumed
    += BATTERY_RX_POWER * (simclock() - accnoise-
>stats.turnOnTime);
```

The total Energy consumed is:

Energy consumption for each node = (energy of transmit that related to number of sending in each node + common energy that related to simulation time) on the other hand, when number of transmit is constant the energy consumption only

ralated to "simulation time".

```
Total energy = BATTERY_RX_POWER * (simclock() -
accnoise- >stats.turnOnTime)+accnoise-
>stats.energyConsumed;
Transmit energy = accnoise->stats.energyConsumed;
```

therefore to compute the Transmission energy only , change the glomosim program in "radio-accnose.pc" file to compute only transmitting energy alone.

8 SIMULATION PARAMETERS

Simulation Parameters	
Network Size	500 X 500 m
Mobility of Nodes	20,40
Range of each Node	625 m
Mobility Model	Random
Minimum Node Speed	5-20 m/sec
Pause Time	0,4,8,16 and 24 sec
Data Rate	One Message per minute
Time	500 seconds

TABLE 2: SIMULATION PARAMETERS

9 SIMULATION RESULTS

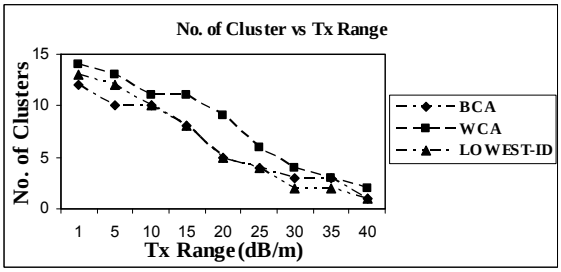


Figure 1: No. of Cluster vs Tx Range

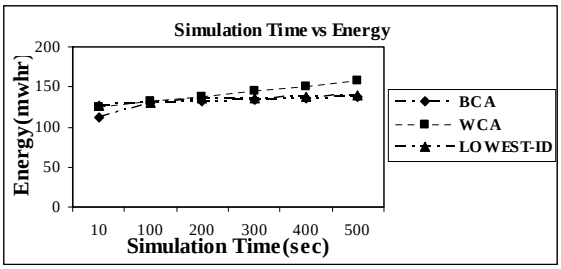


Figure 2: Simulation Time vs Energy

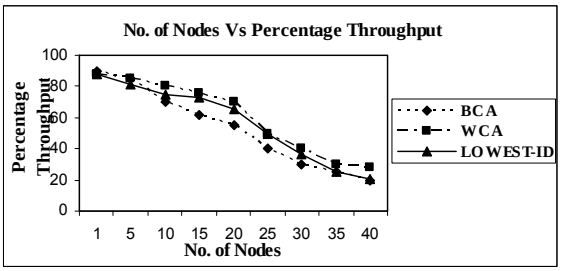


Figure 3: No. of Nodes vs Percentage Throughput

10 CONCLUSION

We have proposed a new clustering algorithm called Cluster Based Algorithm. BCA is an efficient routing protocol for managing the energy usage and security in MANETs. It is a dynamic routing protocol with controlled routing overheads. The routing packets are concentrated in the best paths regions. This allows better optimization with lower number of packets. In the observations we have seen that the above technique gives good performance in some stressful situation like smaller number of nodes and lower load and or mobility. This is an efficient method for managing the energy usage and security in MANETs. This will be the new method to compute stability more simple and possible to be used in ad hoc network to improve the Quality of Service.

REFERENCES

[1] Handbook of wireless networks and mobile computing, Ch. Mobile Ad hoc networks and routing Protocols, Written by YWS, & Edited by Ivan Stojmenovic, University of Ottawa.

[2] L. Bajaj, M. Takai, R. Ahuja, K. Tang, R. Bagrodia, and M. Gerla, "Glomosim: a scalable network simulation environment", Computer Science Department, University of California, Los Angeles, Calif, USA, 1999.

[3] C. E. Perkins, E. M. Royer, S. R. Das, and M. K. Marina, "Performance comparison of two on-demand routing protocols for ad hoc networks," IEEE Pers. Commun., vol. 8, no. 1, pp. 16–28, 2001. T. S. Rappaport, Wireless Communications, Principles and Practice, Prentice-Hall, Upper Saddle River, NJ, USA, 1996.

[4] Boukerche, R. W. Pazzi, and R.B Araujo, Fault-tolerant wireless sensor network routing protocols for the supervision of context-aware physical environments, Journal of Parallel and Distributed Computing, Volume 66, Issue 4, Algorithms for Wireless and Ad-Hoc Networks, April 2006, 586- 599.

[5] D. Estrin, D. Culler, K. Pister, and G. Sukhatme. Connecting the physical world with pervasive networks. IEEE Pervasive Computing, pages 59 – 69, January- March 2002.

[6] Luo J., Hubaux J-P., "Joint Mobility and Routing for Lifetime Elongation in Wireless Sensor Networks", INFOCOM 2005. 24th Annual Joint Conference of the IEEE Computer and Communications Societies, pages 1735-1746, Miami, March 2005.

[7] D. Ganesan, R. Govindan, S. Shenker, and D. Estrin, "Highly-resilient, Energy-efficient multipath Routing in Wireless Sensor Networks" Mobile Computing and Communications Review, vol. 4, no. 5, October 2008

[8] Johnson David B., Routing in Ad hoc networks of mobile hosts, proceeding of IEEE workshop on mobile computing system and applications, December 1994.

[9] Handbook of wireless networks and mobile computing, Ch. Mobile Ad hoc networks, Written by Silvia Giordano, ICCA, Switzerland, & Edited by Ivan Stojmenovic, University of Ottawa,

[10] W. Yu and J. Lee, "DSR-based energy-aware routing protocols in ad hoc networks," in Proc. ICWN Conference, June 2002.

[11] J. Broch, D.B. Johnson and D.A Maltz : The Dynamic Source Routing Protocol for Mobile Ad-Hoc Networks, IETF Internet Draft, draft-ietf-manet-dsr-01.txt, December 1998.

[12] Perkins C. Ad Hoc Networking: Addison-Wesley: 2001.

[13] Barbeau M., Kranakis E., Krizanc D., Morin P., "Improving Distance Based Geographic Location Techniques in Sensor Networks", In 3rd International Conference on ADHOC Networks and Wireless. Vancouver, British Columbia, July 2004.

[14] X. Masip- Bruin, M. Yannuzzi, J Domingo Pascal, A. Fonte, M. Curada, E. Monterio, F. Kuipers, P. Van Mieghem, S. Avallone, G. Ventre, P. Arnada- Gutierrez, M. Hillich, R. Steinmetz, L Iannone, K. Salamatian, Research Challenges in QoS routing, Computer Communications 29 (2006) 563-581.

[15] C. Siva Ram Murthy and B. S. Manoj, Quality of Service in Ad hoc Wireless Networks, chapter 10 in Ad hoc Wireless Networks , Architecture and Protocols edited by Printice hall communication s engineering and emerging technologies series, Theodore S. Rappaport, Series, pp 505-583.

[16] Imrich chlamtac, Marco Conti, Jennifer J. N. Liu, Mobile Adhoc Network: Imperative and Challenges, Ad hoc Networks 1 (2003) 13-64.

[17] Edited by Ivan Stojmenovic, Mobile Adhoc Networks, chapter 15 in handbook of wireless networks and mobile computing, by Willy Interscience, pp 325-346.

[18] Edited by Ivan Stojmenovic, Security and Fraud detection in Mobile and Wireless Network, chapter 14 in handbook of wireless networks and mobile computing, by Willy Interscience, pp 309-322.

[19] Dimitris Vassis, Georgios Kormentzas, erformance analysis of IEEE 802.11 ad hoc networks in the presence of exposed terminals, Ad hoc networks, volume , issue 3, May 2008 pp. 474-482

[20] Dzmitry Kliazovich, Fabrizio Granelli, Crosslayer congestion control in ad hoc wireless networks, volume 4, issue 6, November 2006, pp 687-708 *2009*

AUTHORS PROFILE

Prof. M.N. Doja is currently the professor and head in the Department of Computer Engineering and Founder Head of the Department, F/o Engineering & Technology in Jamia Millia Islamia (Central University), New Delhi. Dr. Doja research interests includes Fuzzy Systems, computer networks, Internet and mobile computing and Mobile Ad hoc Networks. He has the 22 years of research experience. He has published more than 100 research papers in National and International Journals.

Mohd. Amjad is currently working as Assistant Professor in the Department of Computer Engineering, F/o Engineering & Technology, Jamia Millia Islamia (Central University), New Delhi. He received B.Tech. degree from A.M.U. Aligarh in computer Engineering and M.Tech. degree in Information Technology from GGSIP University New Delhi. He is currently a Ph.D. scholar in the Department of Computer Engg. Jamia Millia Islamia. His research interests includes Network Security, Internet and mobile computing, Mobile Ad hoc Networks and wireless sensor networks.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

IJCSIS REVIEWERS' LIST

Assist Prof (Dr.) M. Emre Celebi, Louisiana State University in Shreveport, USA
Dr. Lam Hong Lee, Universiti Tunku Abdul Rahman, Malaysia
Dr. Shimon K. Modi, Director of Research BSPA Labs, Purdue University, USA
Dr. Jianguo Ding, Norwegian University of Science and Technology (NTNU), Norway
Assoc. Prof. N. Jaisankar, VIT University, Vellore,Tamilnadu, India
Dr. Amogh Kavimandan, The Mathworks Inc., USA
Dr. Ramasamy Mariappan, Vinayaka Missions University, India
Dr. Yong Li, School of Electronic and Information Engineering, Beijing Jiaotong University, P.R. China
Assist. Prof. Sugam Sharma, NIET, India / Iowa State University, USA
Dr. Jorge A. Ruiz-Vanoye, Universidad Autónoma del Estado de Morelos, Mexico
Dr. Neeraj Kumar, SMVD University, Katra (J&K), India
Dr Genge Bela, "Petru Maior" University of Targu Mures, Romania
Dr. Junjie Peng, Shanghai University, P. R. China
Dr. Ilhem LENGLIZ, HANA Group - CRISTAL Laboratory, Tunisia
Prof. Dr. Durgesh Kumar Mishra, Acropolis Institute of Technology and Research, Indore, MP, India
Jorge L. Hernández-Ardieta, University Carlos III of Madrid, Spain
Prof. Dr.C.Suresh Gnana Dhas, Anna University, India
Mrs Li Fang, Nanyang Technological University, Singapore
Prof. Pijush Biswas, RCC Institute of Information Technology, India
Dr. Siddhivinayak Kulkarni, University of Ballarat, Ballarat, Victoria, Australia
Dr. A. Arul Lawrence, Royal College of Engineering & Technology, India
Mr. Wongyos Keardsri, Chulalongkorn University, Bangkok, Thailand
Mr. Somesh Kumar Dewangan, CSVTU Bhilai (C.G.)/ Dimat Raipur, India
Mr. Hayder N. Jasem, University Putra Malaysia, Malaysia
Mr. A.V.Senthil Kumar, C. M. S. College of Science and Commerce, India
Mr. R. S. Karthik, C. M. S. College of Science and Commerce, India
Mr. P. Vasant, University Technology Petronas, Malaysia
Mr. Wong Kok Seng, Soongsil University, Seoul, South Korea
Mr. Praveen Ranjan Srivastava, BITS PILANI, India
Mr. Kong Sang Kelvin, Leong, The Hong Kong Polytechnic University, Hong Kong
Mr. Mohd Nazri Ismail, Universiti Kuala Lumpur, Malaysia
Dr. Rami J. Matarneh, Al-isra Private University, Amman, Jordan
Dr Ojesanmi Olusegun Ayodeji, Ajayi Crowther University, Oyo, Nigeria
Dr. Riktesh Srivastava, Skyline University, UAE
Dr. Oras F. Baker, UCSI University - Kuala Lumpur, Malaysia
Dr. Ahmed S. Ghiduk, Faculty of Science, Beni-Suef University, Egypt
and Department of Computer science, Taif University, Saudi Arabia
Mr. Tirthankar Gayen, IIT Kharagpur, India
Ms. Huei-Ru Tseng, National Chiao Tung University, Taiwan

CALL FOR PAPERS
International Journal of Computer Science and Information Security
January - December
IJCSIS 2012
ISSN: 1947-5500
http://sites.google.com/site/ijcsis/

International Journal Computer Science and Information Security, IJCSIS, is the premier scholarly venue in the areas of computer science and security issues. IJCSIS 2011 will provide a high profile, leading edge platform for researchers and engineers alike to publish state-of-the-art research in the respective fields of information technology and communication security. The journal will feature a diverse mixture of publication articles including core and applied computer science related topics.

Authors are solicited to contribute to the special issue by submitting articles that illustrate research results, projects, surveying works and industrial experiences that describe significant advances in the following areas, but are not limited to. Submissions may span a broad range of topics, e.g.:

Track A: Security

Access control, Anonymity, Audit and audit reduction & Authentication and authorization, Applied cryptography, Cryptanalysis, Digital Signatures, Biometric security, Boundary control devices, Certification and accreditation, Cross-layer design for security, Security & Network Management, Data and system integrity, Database security, Defensive information warfare, Denial of service protection, Intrusion Detection, Anti-malware, Distributed systems security, Electronic commerce, E-mail security, Spam, Phishing, E-mail fraud, Virus, worms, Trojan Protection, Grid security, Information hiding and watermarking & Information survivability, Insider threat protection, Integrity
Intellectual property protection, Internet/Intranet Security, Key management and key recovery, Language-based security, Mobile and wireless security, Mobile, Ad Hoc and Sensor Network Security, Monitoring and surveillance, Multimedia security ,Operating system security, Peer-to-peer security, Performance Evaluations of Protocols & Security Application, Privacy and data protection, Product evaluation criteria and compliance, Risk evaluation and security certification, Risk/vulnerability assessment, Security & Network Management, Security Models & protocols, Security threats & countermeasures (DDoS, MiM, Session Hijacking, Replay attack etc,), Trusted computing, Ubiquitous Computing Security, Virtualization security, VoIP security, Web 2.0 security, Submission Procedures, Active Defense Systems, Adaptive Defense Systems, Benchmark, Analysis and Evaluation of Security Systems, Distributed Access Control and Trust Management, Distributed Attack Systems and Mechanisms, Distributed Intrusion Detection/Prevention Systems, Denial-of-Service Attacks and Countermeasures, High Performance Security Systems, Identity Management and Authentication, Implementation, Deployment and Management of Security Systems, Intelligent Defense Systems, Internet and Network Forensics, Large-scale Attacks and Defense, RFID Security and Privacy, Security Architectures in Distributed Network Systems, Security for Critical Infrastructures, Security for P2P systems and Grid Systems, Security in E-Commerce, Security and Privacy in Wireless Networks, Secure Mobile Agents and Mobile Code, Security Protocols, Security Simulation and Tools, Security Theory and Tools, Standards and Assurance Methods, Trusted Computing, Viruses, Worms, and Other Malicious Code, World Wide Web Security, Novel and emerging secure architecture, Study of attack strategies, attack modeling, Case studies and analysis of actual attacks, Continuity of Operations during an attack, Key management, Trust management, Intrusion detection techniques, Intrusion response, alarm management, and correlation analysis, Study of tradeoffs between security and system performance, Intrusion tolerance systems, Secure protocols, Security in wireless networks (e.g. mesh networks, sensor networks, etc.), Cryptography and Secure Communications, Computer Forensics, Recovery and Healing, Security Visualization, Formal Methods in Security, Principles for Designing a Secure Computing System, Autonomic Security, Internet Security, Security in Health Care Systems, Security Solutions Using Reconfigurable Computing, Adaptive and Intelligent Defense Systems, Authentication and Access control, Denial of service attacks and countermeasures, Identity, Route and

Location Anonymity schemes, Intrusion detection and prevention techniques, Cryptography, encryption algorithms and Key management schemes, Secure routing schemes, Secure neighbor discovery and localization, Trust establishment and maintenance, Confidentiality and data integrity, Security architectures, deployments and solutions, Emerging threats to cloud-based services, Security model for new services, Cloud-aware web service security, Information hiding in Cloud Computing, Securing distributed data storage in cloud, Security, privacy and trust in mobile computing systems and applications, **Middleware security & Security features:** middleware software is an asset on
its own and has to be protected, interaction between security-specific and other middleware features, e.g., context-awareness, **Middleware-level security monitoring and measurement:** metrics and mechanisms for quantification and evaluation of security enforced by the middleware, **Security co-design:** trade-off and co-design between application-based and middleware-based security, **Policy-based management:** innovative support for policy-based definition and enforcement of security concerns, **Identification and authentication mechanisms:** Means to capture application specific constraints in defining and enforcing access control rules, **Middleware-oriented security patterns:** identification of patterns for sound, reusable security, **Security in aspect-based middleware:** mechanisms for isolating and enforcing security aspects, **Security in agent-based platforms:** protection for mobile code and platforms, Smart Devices: Biometrics, National ID cards, Embedded Systems Security and TPMs, RFID Systems Security, Smart Card Security, Pervasive Systems: Digital Rights Management (DRM) in pervasive environments, Intrusion Detection and Information Filtering, Localization Systems Security (Tracking of People and Goods), Mobile Commerce Security, Privacy Enhancing Technologies, Security Protocols (for Identification and Authentication, Confidentiality and Privacy, and Integrity), Ubiquitous Networks: Ad Hoc Networks Security, Delay-Tolerant Network Security, Domestic Network Security, Peer-to-Peer Networks Security, Security Issues in Mobile and Ubiquitous Networks, Security of GSM/GPRS/UMTS Systems, Sensor Networks Security, Vehicular Network Security, Wireless Communication Security: Bluetooth, NFC, WiFi, WiMAX, WiMedia, others

This Track will emphasize the design, implementation, management and applications of computer communications, networks and services. Topics of mostly theoretical nature are also welcome, provided there is clear practical potential in applying the results of such work.

Track B: Computer Science

Broadband wireless technologies: LTE, WiMAX, WiRAN, HSDPA, HSUPA, Resource allocation and interference management, Quality of service and scheduling methods, Capacity planning and dimensioning, Cross-layer design and Physical layer based issue, Interworking architecture and interoperability, Relay assisted and cooperative communications, Location and provisioning and mobility management, Call admission and flow/congestion control, Performance optimization, Channel capacity modeling and analysis, Middleware Issues: Event-based, publish/subscribe, and message-oriented middleware, Reconfigurable, adaptable, and reflective middleware approaches, Middleware solutions for reliability, fault tolerance, and quality-of-service, Scalability of middleware, Context-aware middleware, Autonomic and self-managing middleware, Evaluation techniques for middleware solutions, Formal methods and tools for designing, verifying, and evaluating, middleware, Software engineering techniques for middleware, Service oriented middleware, Agent-based middleware, Security middleware, Network Applications: Network-based automation, Cloud applications, Ubiquitous and pervasive applications, Collaborative applications, RFID and sensor network applications, Mobile applications, Smart home applications, Infrastructure monitoring and control applications, Remote health monitoring, GPS and location-based applications, Networked vehicles applications, Alert applications, Embeded Computer System, Advanced Control Systems, and Intelligent Control : Advanced control and measurement, computer and microprocessor-based control, signal processing, estimation and identification techniques, application specific IC's, nonlinear and adaptive control, optimal and robot control, intelligent control, evolutionary computing, and intelligent systems, instrumentation subject to critical conditions, automotive, marine and aero-space control and all other control applications, Intelligent Control System, Wiring/Wireless Sensor, Signal Control System. Sensors, Actuators and Systems Integration : Intelligent sensors and actuators, multisensor fusion, sensor array and multi-channel processing, micro/nano technology, microsensors and microactuators, instrumentation electronics, MEMS and system integration, wireless sensor, Network Sensor, Hybrid

Sensor, Distributed Sensor Networks. Signal and Image Processing : Digital signal processing theory, methods, DSP implementation, speech processing, image and multidimensional signal processing, Image analysis and processing, Image and Multimedia applications, Real-time multimedia signal processing, Computer vision, Emerging signal processing areas, Remote Sensing, Signal processing in education. Industrial Informatics: Industrial applications of neural networks, fuzzy algorithms, Neuro-Fuzzy application, bioInformatics, real-time computer control, real-time information systems, human-machine interfaces, CAD/CAM/CAT/CIM, virtual reality, industrial communications, flexible manufacturing systems, industrial automated process, Data Storage Management, Harddisk control, Supply Chain Management, Logistics applications, Power plant automation, Drives automation. Information Technology, Management of Information System : Management information systems, Information Management, Nursing information management, Information System, Information Technology and their application, Data retrieval, Data Base Management, Decision analysis methods, Information processing, Operations research, E-Business, E-Commerce, E-Government, Computer Business, Security and risk management, Medical imaging, Biotechnology, Bio-Medicine, Computer-based information systems in health care, Changing Access to Patient Information, Healthcare Management Information Technology. Communication/Computer Network, Transportation Application : On-board diagnostics, Active safety systems, Communication systems, Wireless technology, Communication application, Navigation and Guidance, Vision-based applications, Speech interface, Sensor fusion, Networking theory and technologies, Transportation information, Autonomous vehicle, Vehicle application of affective computing, Advance Computing technology and their application : Broadband and intelligent networks, Data Mining, Data fusion, Computational intelligence, Information and data security, Information indexing and retrieval, Information processing, Information systems and applications, Internet applications and performances, Knowledge based systems, Knowledge management, Software Engineering, Decision making, Mobile networks and services, Network management and services, Neural Network, Fuzzy logics, Neuro-Fuzzy, Expert approaches, Innovation Technology and Management : Innovation and product development, Emerging advances in business and its applications, Creativity in Internet management and retailing, B2B and B2C management, Electronic transceiver device for Retail Marketing Industries, Facilities planning and management, Innovative pervasive computing applications, Programming paradigms for pervasive systems, Software evolution and maintenance in pervasive systems, Middleware services and agent technologies, Adaptive, autonomic and context-aware computing, Mobile/Wireless computing systems and services in pervasive computing, Energy-efficient and green pervasive computing, Communication architectures for pervasive computing, Ad hoc networks for pervasive communications, Pervasive opportunistic communications and applications, Enabling technologies for pervasive systems (e.g., wireless BAN, PAN), Positioning and tracking technologies, Sensors and RFID in pervasive systems, Multimodal sensing and context for pervasive applications, Pervasive sensing, perception and semantic interpretation, Smart devices and intelligent environments, Trust, security and privacy issues in pervasive systems, User interfaces and interaction models, Virtual immersive communications, Wearable computers, Standards and interfaces for pervasive computing environments, Social and economic models for pervasive systems, Active and Programmable Networks, Ad Hoc & Sensor Network, Congestion and/or Flow Control, Content Distribution, Grid Networking, High-speed Network Architectures, Internet Services and Applications, Optical Networks, Mobile and Wireless Networks, Network Modeling and Simulation, Multicast, Multimedia Communications, Network Control and Management, Network Protocols, Network Performance, Network Measurement, Peer to Peer and Overlay Networks, Quality of Service and Quality of Experience, Ubiquitous Networks, Crosscutting Themes – Internet Technologies, Infrastructure, Services and Applications; Open Source Tools, Open Models and Architectures; Security, Privacy and Trust; Navigation Systems, Location Based Services; Social Networks and Online Communities; ICT Convergence, Digital Economy and Digital Divide, Neural Networks, Pattern Recognition, Computer Vision, Advanced Computing Architectures and New Programming Models, Visualization and Virtual Reality as Applied to Computational Science, Computer Architecture and Embedded Systems, Technology in Education, Theoretical Computer Science, Computing Ethics, Computing Practices & Applications

Authors are invited to submit papers through e-mail ijcsiseditor@gmail.com. Submissions must be original and should not have been published previously or be under consideration for publication while being evaluated by IJCSIS. Before submission authors should carefully read over the journal's Author Guidelines, which are located at http://sites.google.com/site/ijcsis/authors-notes .

ISSN 1947 5500